THE RELEVANCE OF COGNITIVE STYLE AND MOTIVATION TO ACADEMIC AND NON-ACADEMIC ACHIEVEMENT

by

Parween Hasan

Thesis presented for the Degree of Ph.D.

University of Edinburgh

April 1973

ACKNOWLEDGEMENTS

I had been out of touch with the world of academic research for five years when I started work on this thesis with some trepidation. That it can now be presented in a completed form, is, to a great extent, due to the careful and critical guidance given by my supervisor, Dr. A.E.G. Pilliner, who inspite of his many engagements in this country and abroad, has always found time to comment upon and discuss the various drafts submitted to him. I am most grateful to him for his help and advice throughout the time that I was engaged in the research reported in this thesis.

I would also like to thank Professor Liam Hudson for his interest in the progress of my research and for his encouragement towards its completion.

During the early stages of this research, Dr. Boris Semeonoff of the Department of Psychology made very useful suggestions regarding the relevant literature on the topic of achievement motivation, and he also allowed me to use his TAT slides. I am grateful to him for his help in this respect.

I would also like to record my thanks to the Board of Intermediate and Secondary Education, Karachi (Pakistan), for allowing me leave of absence from my job, and to the British Council for offering me a scholarship to enable me to resume my studies at the University of Edinburgh.

It would not have been possible for me to pursue this research without the help and co-operation of the staff of the Local Authority Education Office, the Careers Office and the school where I carried out the testing. I am particularly grateful to Mr. T.D. Black of the Careers Office and to Mr. Sutherland, Deputy Headmaster of Forrester Secondary School for finding time to provide me with the required information and for arranging the testing schedule.

Thanks are also due to Miss Avril Myles who undertook to do the typing from a hand-written draft which, I am sure must have been difficult to decipher at times.

And finally, I would like to express my appreciation for the resourcefulness and understanding with which my husband and six year old son dealt with the ups and downs of living with a researching wife and mother.

CONTENTS

		Page
	Glossary of Terms Used	i-ii
	List of Tables	iii
	Summary	vii
Chapter I	INTRODUCTION	1

SECTION ONE

BACKGROUND OF THE PRESENT STUDY

Chapter II	CONVERGENT AND DISCRIMINANT VALIDITY OF DIVERGENT THINKING TESTS	10
	1. Evidence from Wallach and Kogan's (1965) review.	10
	2. Studies since Wallach's (1970) review	11
	3. The threshold hypothesis	30
	4. The threshold hypothesis and IQ-divergent thinking correlations	33
Chapter III	PREDICTIVE VALIDITY OF DIVERGENT THINKING TESTS	47
	1. Divergent thinking and academic attainment	48
	2. Divergent thinking and non-academic accomplishments	62
Chapter IV	THE ROLE OF MOTIVATION IN DIVERGENT THINKING AND NON-ACADEMIC ACHIEVEMENT	88
	1. Motivation and divergent thinking	90
	2. Motivation and non-academic accomplishment	103

SECTION TWO

DESCRIPTION OF THE PRESENT STUDY

Chapter V	AIMS AND PLAN OF RESEARCH	108
	1. Aims	108
	2. Plan of research	110
	(a) The school	110
	(b) The pupils	111
	(c) Groups studied	111
	(d) Administration of tests and questionnaire	121

		Page
Chapter VI	DESCRIPTION OF TESTS AND QUESTIONNAIRE	126
	1. The TAT	126
	(a) Slides used	126
	(b) Administration	130
	(c) Scoring	131
	2. Divergent Thinking Tests	148
	(a) Description of tests	148
	(b) Administration	149
	(c) Scoring	153
	3. Moray House Intelligence Test	163
	4. Non-academic Accomplishments Questionnaire	163
	(a) Format of questionnaire	163
	(b) Administration	164
	(c) Scoring	165

SECTION THREE

STATEMENT AND DISCUSSION OF RESULTS

Chapter VII	ANALYSIS OF DATA FOR THE WHOLE SAMPLE	169
	1. Distribution of raw scores	169
	2. Correlation between IQ and divergent thinking	170
	3. Inter-correlations among IQ, English, Arithmetic, Career Choice and Socio-economic Background	175
	4. Correlations of divergent thinking with English, Arithmetic, Non-academic Accomplishments, Socio-economic Background and Career Choice	181
	5. Correlations between Motivation Index and other variables	186
	6. Correlations between nAch and other variables	192
	7. Multiple regression analysis	202

		Page
Chapter VIII	ANALYSIS OF RESULTS FOR THREE LEVELS OF IQ	208
	1. Selection of cut-off points to form High, Middle and Low IQ groups	208
	2. IQ-divergent thinking correlations at three levels of IQ	211
	3. Correlation between DT and dependent variables at three levels of IQ	216
	4. Correlations of Motivation Index, nAch and Socio-economic Background with dependent variables	218
Chapter IX	ANALYSIS OF RESULTS FOR SCHOOL CLASSES	220
	1. IQ-divergent thinking correlations for Top, Middle and Leavers' group	221
	2. Correlations between independent and dependent variables for Top, Middle and Leavers' sets	226
Chapter X	ANALYSIS OF RESULTS FOR EXPERIMENTAL GROUPS	229
	1. Analysis of variance results	230
	2. Paired comparisons of three experimental groups	234
	3. Comparison of experimental groups for participation in extra-curricular activities	237
Chapter XI	SUMMARY AND CONCLUSIONS	242
APPENDIX A	: Statistical Tables	252
APPENDIX B	: Tests and questionnaire	263
APPENDIX C	: Kelley's (1947) formulae and computer programme	276
APPENDIX D	: Review of Literature chart	279
REFERENCES		286

GLOSSARY OF TERMS USED

Tests and other variables:

DT	Divergent Thinking Score. Total of Fluency and Uniqueness scores on Uses, Similarities and Consequences tests. Raw scores normalised to a mean of 110 and standard deviation of 11 to make the scores comparable with IQ.
MI	Motivation Index derived from an analysis of stories written to TAT slides. Scoring based on the method of Story Sequence Analysis (Arnold, 1962).
n-Ach	Need Achievement score obtained from the same TAT slides used for MI above, but scoring based on D-2 version of McClelland's system described in Saddaca, Clark and Ricciuti (1957).
NAAQ	Non-academic Accomplishments Questionnaire taken from Wallach and Wing (1969) with necessary modifications to make it suitable for secondary school pupils. Raw scores normalised to a mean of 10 and a standard deviation of 5.
ENG	English attainment score. Teachers' marks in school examination scaled on VRQ for each class separately, using the NFER method described by Yates and Pidgeon (1957).
ARITH	Arithmetic attainment score. Teachers' marks scaled as in ENG.
SEB	Socio-economic background. Father's occupation categorised according to the Registrar General's classification (1970) on a six-point scale.
JOB CHOICE	Job aspiration of pupils also ranked according to the Registrar General's classification.

Grouping of Pupils

School Classes

TOP SET	Classes 4C1 + 3C1
MID SET	Classes 4C4 + 3C4
LEAVERS	Classes 3B2 leaving school at 15.

<u>VRQ levels</u>

HIGH IQ	Pupils from total sample within IQ range of 123-140
MID IQ	Pupils from total sample within IQ range 103-122.
LOW IQ	Pupils from total sample within IQ 85-102.

<u>Experimental Groups</u>

H-H	High-High group. Pupils in the top third of the IQ <u>and</u> DT distributions.
H-IQ	High-IQ group. Pupils in the top third of the IQ distribution but <u>not</u> in the top third of the DT distribution.
H-DT	High-Divergent thinking group. Pupils in the top third of DT but <u>not</u> in the top third of the IQ distribution.
L-L	Low-Low group. Pupils in the lowest third of the IQ <u>and</u> DT distributions.

LIST OF TABLES

No.	Table	Page
II.1	Average Correlations Based on Biggs et al (1971) Data	11
II.2	Average Correlations Based on Cropley's (1972) Data	13
II.3	Average Correlations Based on Kogan and Pankove's (1972) Data	19
II.4	Average Correlations Based on Di Scipio's (1971) Data	20
II.5	Average Correlations from Ward, Kogan and Pankove's (1972) Study	24
III.1	Summary of Wallach and Wing's (1969) Findings Regarding Mean Non-Academic Accomplishments Scores of Groups High and Low on IQ, Ideational Productivity and Ideational Uniqueness	78
V.1	Number and Percentage of Pupils in the Top and Middle Sets Taking "Difficult" Subjects	114
V.2	Number and Percentage of Boys in the Top and Middle Sets Taking "Easy" Subjects	114
V.3	Number and Percentage of Girls in the Top and Middle Sets Taking "Easy" Subjects	114
VI.1	Raw Score Statistics of Flexibility and Uniqueness Scores of Divergent Thinking Test	162
VII.1	Average Correlations of DT Items with Flexibility Total, Originality Total, IQ, English and Arithmetic	174
VII.2	Means, Standard Deviations, F, of MI at Four Levels of IQ, DT, nAch, SEB, English, Arithmetic, NAAQ and Career Choice	187-188

		Page
VII.3(a)	Multiple Regression Analysis to Estimate the Contribution of Independent Variables to English	203
VII.3(b)	Multiple Regression Analysis to Estimate the Contribution of Independent Variables to Arithmetic	203
VII.3(c)	Multiple Regression Analysis to Estimate the Contribution of Independent Variables to NAAQ	204
VII.3(d)	Multiple Regression Analysis to Estimate the Contribution of Independent Variables to Career Choice	204
VII.4	Partial Correlations Between SEB and Dependent Variables, Controlling for IQ	205
VII.5	Partial Correlations Between NAAQ and Independent Variables, Controlling for DT	206
VIII.1	Cut-off Points on IQ Scale to Form High, Middle and Low Groups	210
VIII.2	IQ-DT Correlations at Three Levels of IQ	211
VIII.3	IQ-DT Correlations for Whole Sample and Three Levels of IQ in Lytton and Cotton's (1969) Study	214
VIII.4	IQ-DT Correlations at Three Levels of IQ for Boys and Girls in Bennett's (1973) Study	215
IX.1	IQ-DT Correlations for Top, Middle and Leavers' Group	221
IX.2	IQ, DT Means, Standard Deviations, Correlations for Subgroups from Top and Middle Sets, Matched for IQ	225
X.1	Analysis of Variance for Significance of Differences Among Mean Scores of Four Experimental Groups	230

		Page
X.2	Analysis of Variance for Significance of Differences Among Mean Scores of <u>Three</u> Experimental Groups	233
X.3	Means and Significance of Differences Between Pairs of Means of Three Experimental Groups	234
X.4	Analysis of Variance for Orthogonal Comparisons of English Scores of Three Experimental Groups	235
X.5	Number of Pupils in Three Experimental Groups Reporting Participation in Extra-curricular Activities	239
X.6	Number of Pupils in Three Experimental Groups Reported by Teachers to have Participated in Extra-curricular Activities	239

APPENDIX A

1.1	DATA FROM (A) FIVE SCHOOL CLASSES, (B) WHOLE SAMPLE. MEANS, STANDARD DEVIATIONS, SIGNIFICANCE OF DIFFERENCE BETWEEN CLASS MEANS	252
1.2	MEANS, STANDARD DEVIATIONS, SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS OF BOYS AND GIRLS (A) WHOLE SAMPLE (B) NON-LEAVERS ONLY	254
1.3	MEANS, STANDARD DEVIATIONS, SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS OF (A) LEAVERS/NON-LEAVERS (B) THIRD YEAR/FOURTH YEAR NON-LEAVERS	255
1.4	MEANS, STANDARD DEVIATIONS, SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS OF TOP, MIDDLE, LEAVERS' SETS	256
1.5	MEANS, STANDARD DEVIATIONS, SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS OF HIGH, MIDDLE, LOW IQ GROUPS	257

Page

1.6	MEANS, STANDARD DEVIATIONS OF (A) FOUR EXPERIMENTAL GROUPS (B) REST OF THE SAMPLE	258
2.1	INTER-CORRELATIONS OF INDEPENDENT AND DEPENDENT VARIABLES	260
2.2	MEANS, STANDARD DEVIATIONS, INTER-CORRELATIONS OF NAAQ	260
2.3	MEANS, STANDARD DEVIATIONS, INTER-CORRELATIONS OF DT (ITEMS/SUBTESTS), IQ, ENGLISH AND ARITHMETIC	261
2.4	CORRELATION BETWEEN INDEPENDENT AND DEPENDENT VARIABLES FOR TOP, MIDDLE, LEAVERS' SET	262
2.5	CORRELATION BETWEEN INDEPENDENT AND DEPENDENT VARIABLES FOR HIGH, MIDDLE, LOW LEVELS OF IQ	262

SUMMARY

The main purpose of this thesis has been to examine the evidence for the existence of two cognitive styles or modes of thinking: one measured by conventional intelligence tests, the other by open-ended or divergent thinking tests. A related issue under consideration was the relative contributions of IQ, divergent thinking and achievement motivation to academic and <u>non-academic</u> forms of achievement.

In the Introduction, the salient features of Guilford's "structure of intellect" model, from which the current controversy regarding the convergent-divergent modes of thinking is seen to originate,are outlined.

Section One considers validity studies of divergent thinking tests (most of them subsequent to Wallach's 1970 review), with special reference to convergent-discriminant and predictive validity. Evidence suggested that the required differentiation of divergent thinking tests from IQ occurs if these tests are scored for fluency rather than flexibility. But, the poor predictive validity of these tests (at least for academic performance) throws doubt on the practical and conceptual value of fluency scores. Possible explanations for the poor predictive validity of divergent thinking tests are considered.

Section Two outlines the aims and plans of the research reported in this thesis. It also contains a discussion of the rationale of tests and questionnaires used. In considering the threshold hypothesis

the problem of restriction of range is dealt with by dividing the sample into equal variance sub-groups. It is argued that level of IQ alone may not be the crucial factor in determining the IQ-divergent thinking correlations. Pupils' interpretation of the testing situation and the extent to which they perceive it as being different from usual school-work may also be involved. Evidence for such an interaction was sought by matching pupils from the Top and Middle "sets" for IQ before computing IQ-divergent thinking correlations. If level of IQ is the crucial variable in determining these correlations, then correlations for the two matched groups should be similar, if there is an interaction then inspite of the matching on IQ the pattern of correlations should be different. Finally, four experimental groups (High-High, High-IQ, High-DT and Low-Low) were formed on the basis of their joint standing on IQ and divergent thinking scores, to consider the relative contribution of these variables to academic attainment, non-academic achievement and career aspiration.

Section Three contains the statement and discussion of results. Over the whole range of IQ there was a highly significant positive correlation between IQ and divergent thinking. In a three-level split on the IQ scale, the correlations for the low and high groups were not significant, while that for the middle group was. After matching individuals from the Top and Middle groups on IQ, the trend of correlations was completely reversed: for the Top set there was a positive and significant correlation but for the middle group it was negative and non-significant. These results coupled with the non-significant correlation for the low IQ group in the three level analysis (which is contrary to the prediction of the threshold hypothesis) suggested that to consider IQ level alone as the crucial factor in determining IQ-divergent thinking correlations is an over-simplification.

Multiple regression analysis was used to estimate the relative contributions of IQ, divergent thinking, achievement motivation and socio-economic background to : attainment in English and Arithmetic and questionnaire measures of non-academic achievement and career choice. Results suggested that the contributions of the independent variables are complementary rather than additive.

The differential predictive validity of IQ and divergent thinking for academic and non-academic achievement is also suggested by the analysis of variance results for the four experimental groups. For English, the High-High group had the highest mean score; in Arithmetic the High-High and High-IQ groups did not differ significantly, the High-DT group had a significantly lower mean score than either. But, for the criterion of non-academic achievement the <u>High-DT</u> and High-High group had higher scores than the High-IQ group.

The validity of the non-academic accomplishments questionnaire was checked against information in school record cards and Careers Office records on these groups' participation in extra-curricular activities. A chi-squared test showed that for the High-High and High-DT groups significantly more participation was recorded than for the High-IQ group. Test-retest reliability was high over a ten-month period for this questionnaire.

It is concluded that inspite of a moderate degree of overlap between IQ and divergent thinking scores, these two measures have differential predictive validity if the criterion of achievement is extended to include non-academic as well as academic achievements.

SECTION ONE

BACKGROUND OF THE PRESENT STUDY

CHAPTER I

INTRODUCTION

In retrospect, Guilford's address on "Creativity"* given to the American Psychological Association in 1950 seems to have been one of those "accidents" in the history of a discipline which set off a spate of research and theorizing quite unforeseen at the time when the original ideas are put forward. Although attempts

---

*Guilford's and others' subsequent work in this field is largely based on open-ended tests devised by his team at the University of Southern California. Commenting upon the American literature on "creativity" which has treated scores on these tests as an index of real-life creativeness, Hudson (1966) points out that this is an over-simplification and also misleading, since true creativeness and originality are by no means such a "simple affair". Cronbach (1968) has also criticised Wallach and Kogan (1965) for calling, with "uncertain justification" scores derived from a number of open-ended tests, measures of "creativity". Vernon (1966) too considers these tests as "poor samples of original thought or creative invention in the generally accepted sense". The distinction that Cronbach, Hudson and Vernon point to is certainly an important one and in the present thesis the term "divergent thinking" has been used when referring to performance on open-ended tests. However, at times the word "creativity" seems more appropriate as it avoids what Butcher (1968) has called "clumsy periphrasis", and therefore, it has also been used as a synonym for divergent thinking.

had been made earlier to understand the nature of "giftedness", "creativity" and "originality" by Chassell (1916), Spearman (1923), and Terman et al (1925-1951), these authors had not entertained the possibility of considering these special abilities as being distinguishable from what had come to be known as "intelligence". In fact, Spearman had categorically denied the existence of a separate factor of creativity or imagination:

"No such special creative power exists...That which is usually attributed to such a special imaginative or inventive operation can be simply resolved into a correlate eduction combined with a mere reproduction." (Spearman 1923). The Stanford studies of the gifted carried out by Terman et al had also identified giftedness with high I.Q., and a close look at Laura Chassell's items in her "test of originality" clearly shows that most of them were not very different from conventional intelligence test items.*

Shouksmith (1970) provides a comprehensive historical account of these early attempts under the chapter heading "Traditional Psychological Conceptions Relating to Creativity". He summarises the traditional British view quite succintly, as follows: "It does not deny that creativity can be isolated. It suggests, however, that creativity cannot be regarded as a separate unitary factor, equal in status to, and clearly isolated from the "g" factor reflected in a single intelligence test score. Intelligence, in its noegenetic

---

*Of the twelve different types of items she used, only four may be considered open-ended in format. These were Word Building, Economic Prophecies, the Invention for Sheet Music and Novel Situations.

sense is regarded as an essential component of or prerequisite for, creativity. Creative production depends on the operation of high intelligence along with a number of other factors", (pp.116-117).

Guilford's address referred to above, may therefore be considered to mark the real beginning of creativity research - "a boom in American psychological industry" as Hudson (1966) has called it. In that paper Guilford presented agrand design, wide ranging in scope and considerably optimistic in tone about the possibility of understanding the different aspects of creativity. The broad scope of the ideas discussed can be gathered from the fact that in a few pages of the journal American Psychologist, Guilford managed to pack the following topics:Some definitions and questions, The social importance of creativity, Some general theories of the nature of creativity, Development of creativity, Factorial research design and specific hypotheses considering creative ability. However, from this rather discursive and speculative paper, certain points emerge which have been issues of central concern in the research that has followed.

Creativity or Cognitive Style

Does the term creativity refer only to the most outstanding and distinguished performance of a few individuals in society or is it to be considered as a trait that everyone possesses in a greater or lesser degree? In other words, is it another of those psychological dimensions such as introversion - extraversion or field-independence - field-dependence etc., which determine the style of an individual's personality and along which individuals differ?

Guilford's own answer to this question is a compromise between the two alternatives - ultimately he would like to understand the former kind of creativity. However, there are serious methodological problems involved in it, e.g. the establishment of a criterion and the accidental nature of creative insights. Therefore he considers that there are greater possibilities of observing individual differences in creative performance if we revise our standards, accepting examples of lower degrees of distinction (Guilford 1950, p.445). Practical though it may be, this compromise is not easy to accept for it assumes that the difference between the two types of creativity is only one of degree not kind. This is an assumption which seems especially unwarranted when we consider the factorial design based on the administration of paper and pencil tests and their statistical analysis, as Guilford's proposed scheme for studying creativity. Actually what has happened is that Guilford and his team have become so preoccupied with the construction and analysis of the tests of creativity that they have remained at the "lower degrees of distinction" for most of the time.

A number of outstanding workers in the field (Barron 1963, Getzels and Jackson 1962, Torrance 1966a, 1966b, 1966c, Wallach and Kogan 1965) have kept to Guilford's approach of identifying "creativity" on the basis of open-ended tests, but there are some important exceptions to this approach and these fall into two groups. Prominent amongst the first are Knapp (1963), Mackinnon (1962), Roe (1952, 1953a) and Stein (1956) who have quite specifically isolated some eminent individuals in a particular field before proceeding to study them as a creative group. McClelland's (1962) paper "On the Psycho-dynamics of Creative Physical Scientists" comes under the same approach in so far as he has tried, in the earlier parts, to find a theoretical explanation, in psychodynamic terms, of the source of the eminent physical scientist's creativity. But in the second part of this paper, where he describes a piece of research as supportive evidence for his theory, he moves nearer to the more conventional approach in which ordinary people are studied in the hope of finding explanations for the distinguished achievements of the more creative few*.

In the second group fall most of the British workers, particularly Hudson (1966) and Vernon (1964, 1966) who have emphasised that good performance on open-ended tests is a matter of cognitive style rather than an index of creative potential "in the generally accepted sense". It is significant to note that following the distinction made by the above mentioned authors most British publications in this field have adopted the term "divergent thinking" to refer to performance on open-ended tests. They consider divergent thinking as a mode of cognitive functioning, perhaps with its own personality correlates, but make no claims about its relation to creativity in terms of outstanding or distinguished performance**.

---

*McClelland's (1962) research with the metaphors test was carried out with 17 college freshmen and 35 graduate students and teachers. Nowhere is it stated that this group was in anyway outstanding or distinguished for its achievements.

**See for example, publications by: Bhavnani and Hutt (1972), Di Scipio (1971a, 1971b), Haddon and Lytton (1968, 1971) Hargreaves and Bolton (1972), Hudson (1965), Lytton and Cotton (1969) and Vernon (1971).

The Convergent-Divergent Distinction

Another issue that Guilford raised in his 1950 paper and which has become a topic of much research since, is that of the relationship between I.Q. and divergent thinking abilities measured by performance on open-ended tests. In that paper and subsequent publications, Guilford (1956, 1959, 1967) has argued that the conventional intelligence test is only a limited measure of the variety of human abilities that he, on the basis of his factor-analytic studies, believes exist. According to his "structure of intellect" model presented in the above-mentioned publications, Guilford considers that there are three different bases for classifying intellectual abilities. These are

(a) the kind of mental process or "operation" performed,

(b) the kind of material or "content" on which operations are performed, and

(c) the kind of "products" that emerge as a result of an operation on a certain content.

These three modes of classification are further sub-divided into a three-dimensional model of five "operations" x four "contents" x six "products", yielding a total of 120 cells or factors. Guilford has been criticised by, Eysenck (1967), McNemar (1964), Vernon (1965b) and Wallach (1970) for taking such a piece-meal approach to the study of the human intellect.

Eysenck (1967) points out that "this factorial extension of Thurstone's work has appeared almost as a reductio ad absurdum of the whole approach ----- "the model fails to reproduce the essentially heirarchial nature of the data" and by leaving out this feature from his structure of intellect model, "Guilford has truly cut out the Dane from his production of Hamlet".

Not all the details of this model are relevant to the present discussion, but it is interesting to note that of the five operations or processes (cognition, memory, divergent thinking, convergent thinking and evaluation) outlined by Guilford, the distinction he made between convergent and divergent thinking has become the source of a major controversy in the psychological literature of the last ten years. Of course, the crucial issues involved in this controversy are not entirely new ones, as they can be seen to stem from the earlier discussions of two-factor versus the multi-factor theories of intellectual ability. The special significance of Guilford's work lies in having focused attention on the possibility of the existence of a particular group of abilities known as divergent thinking ability, which stand together factorially, but may be separate from convergent thinking or intelligence as measured by conventional intelligence tests. Most of the items used in open-ended tests of divergent thinking also derive from the hypotheses put forward by Guilford in 1950 regarding the different types of abilities required for successful performance on these tests. He specifically named the factors of sensitivity to problems, fluency, novelty, flexibility, synthethizing/analysing ability, reorganisation or redefinition, handling of complexity, and evaluation, with examples of the types of test that might be used in each instance.

Correlates of Divergent Thinking

Discussing the validity of the tests he had suggested, Guilford remarked that it "calls for correlation of factor measures with practical criteria" but at the same time he pointed out that "Creative productivity in everyday life is undoubtedly dependent upon primary traits other than abilities. Motivational factors (interest and attitudes) as well as temperament factors must be significant contributors" (p.454). Interest he explained as "the person's inclination or urge to engage in some type of activity"; by attitude he meant "the tendency to favour or not to favour some type of object or situation" and temperament he used as the description of a person's general emotional disposition - "his optimism, his moodiness, his self-confidence or his nervouseness" (p.444). Thus, Guilford also touched upon the personality correlates of divergent thinking, almost in passing. But, the distinction implied in his research design between "ability" and other "primary traits" has become increasingly blurred as a result of subsequent research* which has treated the convergent-divergent distinction not only at the cognitive level but as a matter of an overall "style" or "mode" of psychosocial functioning of individuals. Vernon (1964) has strongly argued "for research into the home and leisure and educational backgrounds of creative individuals" thereby implying that motivational and social factors are as important in influencing performance as are the intellectual abilities tapped by tests. Hudson (1968) further emphasised the complexity of the subject by showing that it was not only ability but an individual's perception of his environment and his notions about what behaviour is appropriate in a particular context that determine what he will actually do. We therefore find that the distinction suggested by Guilford between convergent and divergent thinking in his structure of intellect model has been extended far beyond the field of intellectual ability alone.

---

*Barron (1963, 1969) Cattell and Butcher (1968), Dellas and Gaier (1970), Getzels and Jackson (1962), Golann (1963), Hudson (1966, 1968), Maddi (1965), Wallach and Kogan 1965), Wallach and Wing (1969).

Thus, the research following on from Guilford's work has centred round two main questions: Firstly, are the abilities measured by open-ended tests sufficiently different from those measured by conventional intelligence tests to justify considering them independent of measured intelligence; and secondly, are there concomitant personality and motivational correlates of convergent and divergent thinking abilities to further validate the initial distinction made by Guilford in the structure of intellect model?

CHAPTER II

CONVERGENT AND DISCRIMINANT VALIDITY OF DIVERGENT THINKING TESTS

Evidence from Wallach and Kogan's Review

Wallach and Kogan (1965) and Wallach (1970) provide a comprehensive critical review of the literature relevant to the question of convergent and discriminant validity of open-ended tests. In evaluating the major studies in the field, these authors used the procedure of computing average correlations among the open-ended tests, among the intelligence measures and then between the open-ended tests and the intelligence measures. This was done to estimate the degree of coherence and differentiation that existed in the predicted direction. In other words, they addressed themselves to the question of whether or not the open-ended tests used in the studies they reviewed, were testing a unified dimension of divergent thinking that was sufficiently independent of measured intelligence. Their conclusion was strongly negative: "The measures that have been construed as indicators of creativity are not indicators of some single psychological dimension parallel to and distinct from the dimension of general intelligence defined by conventional intelligence test indices. On the basis of this evidence, then, there is questionable warrant for proposing the very conceptualisation which most researchers have proposed: that creativity is not intelligence, and that individual differences in creativity possess the same degree of psychological pervasiveness as individual differences in general intelligence" (Wallach and Kogan 1965).

Studies Since Wallach's Review

Other studies which have appeared since Wallach's review of 1970 point to a more equivocal situation. In a study with Australian schoolchildren, Biggs et al (1971) also found that all variables used in their study to measure convergent and divergent thinking intercorrelated positively. Their measures of convergent thinking were Raven's Matrices, Australian Council for Educational Research (ACER) group verbal intelligence test, a Mathematical concept test and another ACER reading comprehension test called Implied Meaning. Divergent thinking was assessed by the Use of Objects tests and a Hidden Figures Test devised by the authors. Biggs et al report obtaining two orthogonal factors from the test scores by a varimax rotation, and on the basis of these factor scores they identified the high-low groups on convergence and divergence. No information is given on the degree of overall correlation between their convergent and divergent tests, but from the data given, average correlations* were worked out to compare them with figures reported by Wallach (1970) from other studies.

Table II.1: Average Correlations Based on Biggs et al (1971) data.

Variables Correlated	Ave. r
Verbal and Non-verbal IQ	554
Four Convergent Tests	511
Four Convergent and Two Divergent Tests	270
Two Divergent Tests	244

*(a) All average correlations worked out for the present study were done by using Fisher's z transformation method (Guilford 1965 pp. 348-349).

(b) Decimal points and plus signs have been omitted from the above and following tables giving average correlation from other studies.

In the above table, the pattern of highest average correlation among convergent thinking measures, followed by the correlation between convergent and divergent measures and the lowest average correlation among the divergent thinking tests, reported by Wallach (1970) from earlier studies, is found once again. That is to say, in the Biggs et al study the divergent thinking tests have slightly more shared variance with convergent thinking tests than among themselves. No wonder then that when Biggs et al compare teacher ratings of the high convergent and high divergent groups they find that both groups receive high ratings for the same qualities such as "conceptualism" "independence" and "asks questions". This happens in spite of the fact that the groups had been formed on the basis of factorially independent variables.

Cropley (1972) has recently reported a five-year longitudinal study of the validity of creativity tests. He followed Wallach and Wing's (1969) procedure of comparing the predictive validity of open-ended tests and conventional intelligence tests for predicitng achievements in the non-academic field. In this study, Cropley concludes that "there is a substantial longitudinal relationship between scores on creativity tests and real life performance several years later in domains of the kind conventionally acknowledged as involving creativity"(p. 123). For the full sample Cropley reports a canonical correlation of .51 between the various divergent thinking tests and the criteria of non-academic accomplishment in art, drama, literature and music. Adding I.Q. to the predictor variables raises this correlation to .53. As in a number of other studies of this kind, Cropley does not give any figures for the degree of relationship between convergent and divergent measures.

He does provide a table of intercorrelations of the different tests. Working out average correlations from this table yielded the following values:

Table II.2 : Average Correlations Based on Cropley's (1972) data.

	Boys (n = 56)	Girls (n = 55)
IQ-DT average correlation	274	244
DT-DT average correlation	139	149

Again, we notice that the relationship between IQ and divergent thinking is stronger than among the divergent thinking measures themselves. In view of this fact, it is not surprising that adding IQ to the predictor battery makes only a marginal difference to the canonical correlation.

It should be noted that in both the above studies divergent thinking tests were scored either for spontaneous flexibility (Biggs et al) or for originality (Cropley). In his review Wallach (1970) has argued that scoring for fluency rather than spontaneous flexibility yields the most valid measure of divergent thinking, with originality occupying a midway position. But, of the four fluency factors (Word fluency, Associational fluency, Ideational fluency and Expressional fluency) isolated by Guilford and his colleagues, Wallach suggests that Word fluency should be "banished from the divergent thinking category" on the grounds that it has been found to correlate with conventional measures of intelligence. In the originality domain too, he makes a distinction between statistically unique responses and "clever" responses. The former he consider valid for divergent thinking measures, not the latter, again because original responses which are scored for "cleverness" (as in the Plot Titles test) tend to correlate highly with IQ. Thus, the scoring procedure used in the Biggs et al study and the Cropley study were precisely of the type that would be expected to yield higher correlations with IQ. Even the originality measure used by Cropley is not that of a statistically unique response of one of its kind in a whole sample (which is how Wallach and Kogan 1965, and Wallach and Wing 1969 have defined uniqueness), but a weighted score, "the magnitude of weights being inversely related to the frequency with which a particular response was found on the protocols of the whole sample..." (Cropley 1972, p.120).

Wallach's contention that it is the fluency factors (minus Word fluency) which yield the maximum differentiation between IQ and divergent thinking abilities is further supported by some of the other recent studies to be considered here. With two exceptions (Hargreaves and Bolton 1972 and Vernon 1971) the rest of these studies (Di Scipio 1971a, 1971b; Kogan and Pankove 1972; Leith 1972; Ward, Kogan and Pankove 1972) had scored the divergent thinking measures for fluency in terms of number of ideas given and they all obtained low IQ/DT correlations as would be expected on Wallach's argument. However, before going on to discuss these studies in some detail, it should be pointed out that although the results of these two sets of studies (those scored for flexibility/originality and those scored for fluency) lend empirical support to Wallach's prediction regarding the optimum conditions for a desired low IQ/DT relationship, it still leaves the theoretical question of the face and predictive validity of divergent thinking tests open. For example, is the purely statistical coherence of open-ended tests and their clear differentiation from conventional measures of intelligence really the ultimate test of their validity?

Vernon (1964) raised this issue in an early critical evaluation of Guilford's factorial approach to the study of creativity:"Just because a set of tests looks as though it involves creativity and gives lowish correlations with *g* or v or k tests does not mean that it measures what we recognise as creativity in daily life" (p.166). Since then a number of concurrent (Hudson, 1966, 1968; Wallach and Kogan 1965, Wallach and Wing 1969) and longitudinal (Cropley 1972, Kogan and Pankove 1972) validation studies have appeared which lend support to the distinction made by Guilford between convergent and divergent thinking, by showing that there are certain personality and motivational correlates of these modes of thinking. But in spite of these studies the question of the predictive validity of divergent thinking tests for creativity in terms of distinguished, original performance still remains as debatable as it was ten years ago. These and other questions about how far Wallach is justified in making the convergent and discriminant validity of open-ended tests as the most crucial issue in the IQ/creativity debate are discussed in more detail in Chapter VI (p. 155) of this thesis.

Of the recent studies which have looked at the question of IQ/DT relationship using fluency and originality scores for the divergent thinking measures, Leith's (1972) and Kogan and Pankove's (1972) offer some evidence which accords with Wallach's formulations. Leith studied a group of German secondary school children varying in age from 9 to 13 years and under two conditions of testing - Moderate Stress (MS) and Reduced Stress (RS). As measures of divergent thinking Leith employed the Associations, Uses and Similarities tests, and scored them for "number of responses" and "novelty".

Raven's Standard Progressive Matrices test was used for obtaining the intelligence variable. Leith reports the degree of relationship between his open-ended tests and IQ in the form of rank correlation. For each of the three age groups (9, 11, 13) and under both testing conditions the intercorrelations among the open-ended tests were invariably higher and significant ($p<.001$) than the IQ/DT correlations. None of the latter is reported to be significant and two in fact are in the negative direction. Unfortunately Leith does not give any information regarding the intelligence level of the group he studied and therefore the possibility that the lack of a significant relationship between creativity and intelligence was due to the selected nature of his sample, cannot be ruled out. On the basis of the threshold hypothesis which predicts a decrease in IQ-DT correlations as the level of intelligence rises, Leith's finding would be according to expectation if the mean IQ of his group was around 115 - 120, but they would have limited generalizability.

In a five year follow-up study with some of the children in Pankove and Kogan's (1968) original sample, Kogan and Pankove (1972) found a complex pattern of IQ-DT correlations when divergent thinking tests were administered under individual and group testing conditions and the results were analysed for boys and girls separately. For example, in both testing conditions, the IQ-DT correlation for boys was positive and significant in the follow-up study, whereas in the earlier study it was non-significant*.

---

*It is relevant to recall that this is an example of the kind of fluctuation Vernon (1964) had anticipated in commenting on the instability of specific ability factors and the better predictive power of a reliable g and v for adult accomplishments.

For girls, the earlier finding of generally low, non-significant and sometimes negative correlations between IQ and divergent thinking was replicated. Kogan and Pankove accept that "this contradictory pattern of findings for males and females does not readily lend itself to interpretation". They think one possible explanation may be that for girls motivational and personality factors are more important in influencing performance on open-ended tests. They find support for this view from the fact that for girls anxiety* and divergent thinking had been found to correlate positively at the fifth grade level (Pankove and Kogan 1968) when all testing was done individually in a game-like context, but the correlation became significantly negative in the follow-up study under group testing conditions. It remained significantly positive even in the follow-up study in the individual testing context. The implication seems to be that for girls there is a strong interaction of personality variables, such as anxiety, with the context in which they are asked to perform certain tasks.

For boys, the changed pattern of IQ/DT correlations from non-significant at the fifth grade level to significantly positive in the follow-up study, is also explained by Kogan and Pankove in terms of the moderating effects of anxiety. They report finding no significant correlations between measures of anxiety and performance on open-ended tests for their male sample and suggest that boys' performance in the follow-up study may have been under stricter cognitive

---

*As a measure of anxiety, Pankove and Kogan (1968) had used a general-anxiety scale and a test-anxiety scale adapted from Sarason et al (1960).

control,* hence the correspondence in their performance on convergent and divergent tasks. In other words, at the tenth grade level the boys did not let themselves go when attempting the open-ended tests; on the contrary, in the individual testing situation they seem to have been over-cautious as the "near significant"(p.< .10) decrease in their mean scores from the fifth grade to tenth grade shows.

To obtain an answer to the question of the convergent and discriminant validity of the open-ended tests used in Kogan and Pankove's study, average correlations had to be computed from the correlation matrix given by the authors. As it is presented, the matrix includes several correlations which may be artificially inflated due to a part-whole relationship. For instance, besides giving the correlation between the Uses and Pattern Meanings test for number of ideas (Productivity) Kogan and Pankove also include in the matrix the correlation between the Uses test and the Composite of Uses and Pattern Meanings for Productivity. Similarly, after intercorrelating the four tests for uniqueness, they sum the uniqueness scores across the four tests and then correlate this sum with each of the uniqueness tests which make up the composite.

---

* In a study of anxiety among elementary school children, Sarason et al (1960) also found that on their questionnaire measures, boys came out with lower anxiety scores than girls. They attribute this finding to the greater defensive of boys, socialised in a culture in which the admission of fear and anxiety by males is looked down upon and considered unmasculine. Pankove and Kogan (1968) also found that for boys low on the defensiveness scale there was a significant negative correlation with creativity (n =-.33 p .05).

In summarising their findings regarding these intercorrelations, the authors give the range of correlations along with the median value of r for inter-test productivity and uniqueness scores, as well as for within-test productivity-uniqueness correlations. However, it appears that this information is based on the whole matrix rather than on independent correlations only. To compute average correlations for comparison with other studies only the independent intercorrelations among the four open-ended tests and between the open-ended tests and IQ were utilized. The values obtained in this way are given below:

Table II.3 : Average Correlations Based on Kogan and Pankove's (1972) data.

	GIRLS		BOYS		TOTAL
	Indiv. n=13	Group. n=34	Indiv. n=16	Group. n=38	n=101
Productivity	790	544	635	670	670
Productivity - IQ	-226	104	438	421	188
Uniqueness	605	350	160	340	340
Uniqueness - IQ	-250	110	230	350	110

It will be noticed that inspite of the different pattern of correlations for the two sexes discussed in the text above, the over-all trend with two exceptions (Uniqueness and Uniqueness - IQ for boys under both testing conditions) is that of higher average correlation among the open-ended tests, when scored for number or uniqueness and lower average correlation between IQ and the open-ended tests. This trend is further confirmed in the last column of the above table, thus lending support to Wallach's (1970) idea that if open-ended tests are scored for number and uniqueness of ideas there will be maximum coherence among them and also a clear differentiation from IQ.

In a cross-cultural study with American and English college women, Di Scipio (1971a) found a similar pattern of IQ-DT and DT-DT correlation as Kogan and Pankove above. As divergent thinking variables, Di Scipio used a Hypotheses Test (Shapiro 1966), Uses Test (Hudson 1966) and an Essay Test. They were scored for the number of hypotheses given, number of uses suggested and number of words used,respectively. Thus the divergent thinking score was based essentially on fluency of production although the Essay Test score seems to be specifically a Word fluency score rather than ideational fluency of the type Wallach (1970) considers most appropriate for divergent thinking. Not surprisingly, therefore, the Essay Test did not correlate significantly with the Uses Test in either the American or the English sample and the correlation with the Hypotheses Test, though significant (.45 and .48 for the American and English groups respectively) was low in comparison with that between the Hypotheses and the Uses Tests (.72, $p<.01$). However, when these inter-correlations are compared with the correlations with IQ the following pattern emerges:

Table II.4 : Average Correlations Based on Di Scipio's (1971a) data.

Variables Correlated	American n=30	English n=30
Divergent Thinking Tests	446	470
IQ - Divergent Thinking Tests	159	119

It should be mentioned that Di Scipio's study discussed above was not concerned directly with examining the relationship between

IQ and divergent thinking as such. He was more interested in a cross-cultural study of the personality correlates of divergent thinking, especially the dimensions of Extraversion-Introversion and Neuroticism-Stability as measured by the Eysenck Personality Inventory, and therefore the information reported above was really incidental. However, a recent British study by Hargreaves and Bolton (1972) undertook to examine the question of: "whether "creativity" is a unitary dimension across and between tests and what relationship this range of abilities bears to IQ" (p.451). They used fifteen divergent and non-divergent tests in this study. The divergent battery included tests such as Consequences, Uses for Things, Picture Meanings, Stories, Picture Completion, Drawing, Word Meanings and Similarities etc. These tests were scored according to the Minnesota protocols (Yamamoto 1964b) for Fluency, Flexibility, Originality and Elaboration. IQ was obtained on Morrisby's General Ability Tests, Verbal and Perceptual. Other non-divergent measures included a test based on Mednick's (1962) Remote Associates Test (RAT), a personality questionnaire and an images test. For the present discussion the most relevant aspect of this study is the finding about the intercorrelations among the divergent tests and their correlation with IQ. Besides this, the relationship between RAT and IQ and RAT and the divergent tests is also of some interest, as this test has often been used as a divergent test in a number of earlier studies (Cropley 1966, Ginsburg and Whittemore 1968 and Hasan and Butcher 1966). In all these studies RAT showed a significant positive correlation with IQ and the magnitude of RAT-IQ correlation was not much lower than the correlation of RAT with other divergent thinking tests.

Hargreaves and Bolton also found that this test correlated significantly with IQ ($r = .53$, $p < .001$) and with the other divergent tests ($r = .42$, $p < .001$) thus lending support to Jackson and Messick's (1965) view that it can hardly be considered an open-ended test.

Hargreaves and Bolton summarised their main findings as follows: "Factor analysis of the resulting intercorrelation led to the conclusion that "creativity" implies an integrated range of abilities represented by the divergent tests which, although related to general intelligence in subjects of average IQ, remains factorially distinct from it; within tests, Fluency, Flexibility, Originality and Elaboration subscores were highly intercorrelated." Although Hargreaves and Bolton do not make a distinction between word fluency and other kinds of fluency as Wallach (1970) does, they strongly reinforce his argument that fluency scores are really "paradigmatic" for divergent thinking. For the whole group of 117, eleven-year old boys and girls, the authors report an average correlation of .55 among divergent tests and .45 between the divergent tests and IQ. It will be recalled that the divergent thinking tests in this study were scored not just for fluency but also for flexibility, originality and elaboration. Perhaps that is why the difference between the average divergent test correlations and the IQ-DT correlation is not as much in this study as in Leith's (1972), Kogan and Pankove's (1972) and Di Scipio's (1971a) studies, in which divergent thinking tests were scored for the number of ideas and (with the exception of Di Scipio's study) for uniqueness.

In an earlier study Cline, Richards and Abe (1962) had also found that the correlation between IQ and DT was lower when the latter was scored for total number of ideas and it was higher when scoring was done for change of category or spontaneous flexibility.

One more piece of research which is in accordance with Wallach's (1970) prediction regarding maximum coherence among divergent thinking tests and maximum differentiation between IQ and divergent thinking, if the latter is scored for number of ideas, is that by Ward, Kogan and Pankove (1972). In a study of the "Incentive Effects on Children's Creativity" these authors also provide some information on the extent to which the divergent thinking tests hold together and may be considered different from conventional IQ. The tests of divergent thinking used in this study were derived from Wallach and Kogan's (1965) work; they were two verbal tests (Uses and Similarities) and two non-verbal (Pattern Meanings and Line Meanings). Scoring was done for quantity <u>and</u> quality. Quantity was defined as "number of ideas given by a child except those few judged to be repetitious or inappropriate". High interjudge reliability is reported for this score ($r = .94$). The quality score was not one of uniqueness in the sense of a single response of its kind for the whole sample, as had been employed by Wallach and Kogan (1965), Wallach and Wing (1969) and others. Instead, it was a score on a seven point scale giving credit for the presence of such qualities as appropriateness or cleverness of response. Interjudge agreement on this score ranged from .74 to .87 with a median r of .78. Inspite of these high interjudge reliabilities, the findings regarding divergent test intercorrelations and IQ-DT correlations were mixed, as the following table shows:

Table II.5 : Average correlations from Ward, Kogan and Pankove's (1972) study

Variables Correlated	Average r (n=191)
Number of ideas score	.59
Number of ideas - IQ	.12*
Quality of ideas score	.22
Quality of ideas - IQ	.21*
Quality of ideas - Number of ideas	.04

*Ward et al report that IQ's were available for only two-thirds of the sample, so presumably these average correlations are based on n < 191.

It is interesting to note that although scoring for number of ideas yields a pattern of correlations in the expected direction, i.e. high intercorrelations among divergent thinking tests and low correlation with IQ; quality of ideas score shows quite a different trend. Not only are the intercorrelations and correlations with IQ almost the same, the quality of ideas score is unrelated to the number of ideas score, thus suggesting that whatever is being tested by this quality score is more akin to conventional intelligence than to divergent thinking. The measure of intelligence used was the Kuhlman-Anderson IQ, from school records, with a mean of 94 and standard deviation of 11.5 for the two-thirds of the sample for whom IQ's were available. In the light of the threshold hypothesis, it is significant to find that inspite of this low average IQ, the prediction of high divergent test intercorrelations and very low IQ-DT correlation, was confirmed for only the number score.

Another finding of this study is relevant to the question of when and how the most valid divergent thinking scores may be obtained and so it ought to be mentioned here. Ward et al report that "when subjects are rewarded for each acceptable idea, they show a general increase in the number of ideas given..." (p. 675). The reward used in the study was a U.S. cent per acceptable idea under "Immediate Reward" and "Delayed Reward" conditions. Besides having to give as many ideas as they could, subjects in the experimental groups were also asked to give as many "good" ideas as they could think of. "Good" ideas were not defined and the children were told that their ideas would be compared with those given by other children in the school. Clearly, these testing contexts have a considerable element of evaluative and competitive pressure, yet the authors found that "Both Immediate and Delayed Reward resulted in increased number of ideas relative to the Control group" (p. 673). Flescher (1963) has also reported a significant positive correlation of .27 between his ideational fluency scores and the General Anxiety Scale, thus showing that anxiety may in fact be related to higher divergent thinking scores.

It will be recalled that Wallach and Kogan (1965) had argued that earlier studies had failed to find a unified dimension of creativity which was sufficiently independent of conventional intelligence because these studies had attempted to assess creativity in a evaluative, stressful context not very different from the usual intelligence testing procedures.

As an alternative, they had suggested that "creativity.... if it is to reveal itself most clearly, requires a frame of reference which is relatively free from the coercion of time limits and relatively free from the stress of knowing that one's behaviour is under close evaluation" (p. 24). Their own research confirmed the effectiveness of this alternative procedure and for sometime it came to be accepted that a permissive, non-evaluative and gamelike testing context was essential for obtaining valid scores on open-ended tests. The main criterion of validity chosen by Wallach and Kogan was that of statistical coherence among the open-ended tests and their unrelatedness to conventional intelligence. However, in view of the complexity of factors which may affect actual performance, Hudson (1968) pointed out that "concern with such clusterings is simplistic" (p.125) and a number of studies have indeed reported rather equivocal or contrary findings regarding the effectiveness of a non-evaluative setting for creativity testing; Kogan and Morgan (1969) for example report: "no clear-cut superiority for test- or game-like contexts in reference to creativity level was observed. Rather, effects varied dependent upon the task, the criterion of creativity, anxiety and defensiveness level of subjects, and the sex of the subject" (p.125). Similarly in the study referred to earlier, Kogan and Pankove (1972) also found that in an individual, face-to-face, non-evaluative testing context, the performance of adolescent boys on open-ended tests declined.

That the whole issue of the effects of testing context and scoring procedures on divergent thinking test scores is much more complex than Wallach and Kogan (1965) had anticipated, is further shown in a study by Vernon (1971), who looked specifically at the question of how the context and scoring variables influence performance. Vernon studied about 400 Canadian boys and girls at the 8th grade level. Half of these pupils were given seven divergent thinking tests under "ordinary, test-like conditions", the other half "under more relaxed and informal, and relatively untimed conditions". The divergent test battery consisted of the now well-known open-ended tests such as Circles, Pattern Meanings, Alternate Uses, Improvements, Similarities etc. These were scored for Unusualness or U responses according to a weighting scheme devised by the author. For purposes of comparison the total number of F responses were also recorded and the U and F scores were found to be highly correlated ($r = .758$). Other variables such as verbal and non-verbal Intelligence, Multiple Vocabulary, Rorschach, Art Interest and Teacher Ratings for Curiosity were also correlated with F and U scores separately. In every case, U scores had higher correlations with all these variables than F scores. Hence, according to Vernon, "in all subsequent work, the U score alone was employed" (p.253). It is difficult to see why simply higher correlations with other variables should make U scores more appropriate. On Wallach's (1970) argument, the fact that they have higher correlations with verbal and non-verbal intelligence than F scores, should make them less valid as an index of divergent thinking. Vernon also concludes that "scores based on grading unusual responses are more consistent" (p.245), but if intertest correlations reported for F and U separately are to be taken as an index of internal consistency of scores, the mean intercorrelation for F is higher than for U, although the difference is small (F mean $r = .338$; U mean $r = .316$).

Vernon himself agrees that "there is little to choose between F and U in consistency over different tests ..." (p. 252). As mentioned before, the correlation between F and U scores is quite high and the difference in consistency negligible, therefore the main findings of Vernon's study would probably have been the same had F scores been used. The choice of scores is mentioned in some detail here as it is one more instance of researcher bias in selecting what he considers to be significant variables.

Besides the conclusion regarding greater consistency of U scores mentioned above, Vernon summarised his other findings as follows:

"(b) Larger numbers of high scores were obtained under relaxed conditions;

(c) The factorial structure of "formal" and "relaxed" scores is generally similar;

(d) Correlations of relaxed scores with other variables, including intelligence measures, are generally higher, and particularly with variables relating to creativity" (p. 245).

One interesting aspect of Vernon's study ought to be mentioned here as it throws some light on an important possible source of variation in performance on open-ended tests. It will be recalled that commenting on the complexity of the interacting variables which influence performance, Hudson (1968) had pointed out that "tests themselves may interact - not only over time, as in the test-retest situation, but within a single testing session.

Individuals may let themselves go on one open-ended test and have no energy or inspiration left for others" (p. 125). Vernon's study provides some data which show something of this kind happening. In Table 4 (p. 253) Vernon gives the raw scores of his seven divergent thinking tests administred under the formal and relaxed conditions, for three percentile levels (90th, 50th and 10th). The table shows that for three of the tests (Pattern Meanings, Uses and Similarities) scores under the relaxed condition are consistently higher at the 90th and 50th percentile, whereas for the other four tests (Circles, Improvements, Topics and Consequences) the differences are either negligible or in the opposite direction. Therefore, it does seem that certain tests appealed to the subjects more than others, and, as Vernon remarks, under the relaxed condition they could apportion their own time to doing them, thus getting better scores on these tests than the formal group and contributing to the higher total score of the relaxed group.

From the review of studies above, no clear answer regarding the convergent and discriminant validity of divergent thinking tests amerges, but, considered together these studies do highlight the complex nature of the variables which have been shown to affect performance on open-ended tests. For example, Wallach and Kogan's (1965) view that divergent thinking scores which are sufficiently differentiated from IQ can be obtained best in a non-evaluative, gamelike testing context now needs to be qualified in the light of the findings of Kogan and Morgan(1969), Kogan and Pankove (1972), Vernon (1972) and Ward et al (1972), regarding the moderating effects of anxiety, age and sex of subjects, the appeal of particular tests and of concrete rewards as an incentive for good performance.

The correlational analysis in the studies discussed above also shows that the pattern of mean intercorrelations among divergent thinking tests and between IQ and divergent thinking tests varies depending mainly on the scoring scheme used. The issue is further complicated by the diversity of tests used to measure divergent thinking, the age range of subjects and the general practice of using IQ's taken from school records and based on tests administrated at an unspecified time before the administration of divergent thinking tests. To summarise the differences along these lines a table of the relevant studies appearing in journal articles has been prepared condensing as much important information as is possible in this way.* Research reports of book-length such as those by Getzels and Jackson (1962), Hudson (1966, 1968) and Wallach and Kogan (1965) have not been included for the obvious reason that they are too long to summarise and are also so well-known by now that it is unnecessary to include them.

The Threshold Hypothesis

One explanation of these conflicting findings regarding the independence of divergent thinking from IQ has been in terms of the "threshold hypothesis" or the "ability gradient theory". According to this explanation, up to a certain level of IQ, performance on open-ended tests has a linear relationship with measured intelligence but beyond this level (a) the relationship between the two breaks down and (b) divergent thinking abilities then begin to make an independent contribution to actual achievement on certain criterion variables such as academic attainment or achievement in the non-academic spheres of activity.

*Appendix D.

Originally, the idea of an "ability gradient" was put forward by J.E. Anderson (1960), not with particular reference to the IQ/DT relationship but in a discussion of the inter-relations of different human abilities, especially the relationship between measured intelligence and performance or "output" in life. He maintained that we can "think of ability levels in terms of thresholds and ask questions as to the amount necessary to carry on a task and then consider the factors that determine function beyond this threshold. There are cut-off points or levels above which the demonstration of ability in relation to environmental demands is determined by the presence of other factors" (p. 25).

It was Torrance (1960, 1962) who first invoked the threshold hypothesis with special reference to IQ/DT relationsip to explain the discrepant findings of different replications of Getzels and Jackson's (1962) results regarding the independence of divergent thinking from IQ. In fact, Torrance went beyond the question of IQ/DT relationships as such and argued that beyond an IQ of 115 to 120, not only does the relationship become weak but also that divergent thinking abilities make a significant contribution to achievement, independently of IQ. Yamamoto (1960,1964b) who worked with Torrance on these early replications also explains the two discrepant replications in terms of the threshold hypothesis as follows: "In both cases, the mean IQ for the general student body was about 100 in contrast to the remaining six cases where the mean was at least one standard deviation higher than this and, in addition, these two schools appeared to have emphasised traditional kinds of learning rather than learning with due emphasis on divergent thinking and activities" (Yamamoto 1964b).

Thus we have reference here not only to Anderson's cut-off levels but to the "other factors" also which he suggested might affect performance.

Others have used the threshold theory not explicitly but by implication to explain the weak relationship, generally reported, between IQ and distinguished achievement. For example, on the basis of his studies at the Institute of Personality Assessment and Research, Barron (1969) repeats his earlier conclusion that "for certain intrinsically creative activities a specifiable minimum IQ is probably necessary in order to engage in the activity at all, but that beyond the minimum, which often is surprisingly low, creativity has little correlation with scores on IQ tests." It is noteworthy that in an earlier publication Barron (1963) had set this "minimum" around IQ 120. Discussing the relationship between intelligence and creativity among writers, Barron (1963) had suggested that: "Over the total range of intelligence and creativity, a low positive correlation, probably in the neighbourhood of .40, obtains; beyond an IQ of 120, however, measured intelligence is a negligible factor in creativity, and the motivational and stylistic variables upon which our research has laid such stress are the major determiners of creativity" (p. 242).

Reporting on his research with "highly creative" architects at the same Institute as Barron, MacKinnon (1962) also concluded that "above a certain required minimum level of intelligence which varies from field to field and in some instances may be surprisingly low, being more intelligent does not guarantee a corresponding increase in creativeness".

Hudson (1966) came to the same conclusion in his research with clever English schoolboys: "The relation of IQ to intellectual distinction seems, in fact, highly complex. As far as one can tell, the relation at low levels of IQ holds quite well. Higher up, however, it dwindles; and above a certain point, a high IQ is of little advantage." McNemar (1964) makes a similar observation on theoretical grounds. Although his statement is rather long, it is worth quoting as it does make the assumptions in his argument quite explicit: "If we have honest to goodness criterion measures of literary or architectural or scientific creativity, the scatter diagram between IQ and such creativity (not normalised, since it makes sense to expect a skewed distribution for actual creativity) will be triangular in shape for unselected cases. That is, at the high IQ levels there will be a very wide range of creativity whereas as we go down to average IQ, and on down to lower levels, the scatter for creativity will be less and less. Having a high IQ is not a guarantee for being creative; having a low IQ means creativity is impossible" (p.789). Thus McNemar brings into the discussion the importance of a sound criterion, a point often overlooked by most researchers in this field. The review of predictive validity studies in the next chapter does show how the value of correlation fluctuates depending on the criterion used.

The Threshold Hypothesis and IQ-DT Relationship

In the discussion of the threshold hypothesis above, there seem to be two related questions involved. First is the question of the relationship between IQ and divergent thinking or any other ability at different levels of IQ, and the second question is about the independent contribution of other cognitive and motivational factors to performance beyond a certain threshold level of IQ.

The studies which have specifically looked at the IQ-DT relationship have not always made this distinction clear, but in the discussion which follows the two aspects are dealt with separately where possible.

Some of the earliest studies to consider the question of IQ-DT relationship with specific reference to the threshold hypothesis have been carried out by Torrance and his colleagues at the Bureau of Educational Research, University of Minnesota. Yamamoto (1964b) reports two of these studies in some detail, one with secondary school pupils, the other at the elementary school level. The secondary school group consisted of 272 boys and girls from grade 9 to 12 at the University of Minnesota High School. The Lorge-Thorndike verbal battery was used as a measure of intelligence; mean IQ of the group studied was 118.32, the standard deviation 15.00 (the test is normed to a mean of 100 and a standard deviation of 16). Two tests from the Minnesota Tests of Creative Thinking were used as divergent thinking measures, and scored for different dimensions such as Fluency, Flexibility, Cleverness, Constructiveness and Adequacy. On the basis of a creativity total obtained in this way a group of pupils scoring in the top 20% was identified for a study of the threshold hypothesis. Thus there were altogether 54 pupils on whom the main study is based. These 54 were further divided into three sub-groups of High IQ (>135), Middle IQ ($<135 >120$) and Low IQ (<120). The Iowa Tests of Educational Development were used to obtain the dependent variable of school achievement. In view of the threshold hypothesis, the main interest of the study was to find out if the school achievement of the groups differed significantly at the three levels of IQ or whether there was an IQ threshold beyond which increase in IQ did not necessarily mean an increase in achievement too.

On the basis of an analysis of variance with the three groups separately as well as with pairs of contrasting groups, Yamamoto concludes that "the "threshold" phenomenon was observed in that the High and Middle Groups achieved significantly better on the Iowa Tests of Educational Development than did the Low Group. The High and Middle Groups were not significantly different in this regard" (p.404). He did not find confirmation of the threshold of intelligence hypothesis in his study with elementary school pupils. For this group there was a significant difference in the achievement scores of the three IQ subgroups, scores on the two variables (IQ and DT) showing a linear increase. Yamamoto also reports a correlation of .30 ($p < .01$) between the creativity total and IQ for the secondary school pupils and of .14 ($p < .01$) for the elementary school.

With reference to the first aspect of the threshold hypothesis mentioned earlier (i.e. lack of any significant relationship between IQ and DT around a certain IQ level) it should be noted that in Yamamoto's study there was a significant correlation between these variables inspite of the fact that the mean IQ of his sample was more than one standard deviation above the population mean and almost approaches the IQ of 120 which has been suggested as the most likely cut-off point by most writers. Therefore, if we are considering this first aspect of the threshold theory it cannot be considered confirmed on the basis of the correlational analysis. As for the second question, that of the added contribution of divergent thinking to achievement beyond a certain IQ level the position still seems equivocal, because although the study does show that beyond the IQ level of 120 increase in IQ is not related to an increase in achievement, it does <u>not</u> show that the high creativity of the subjects (who after all are from the top 20% on this variable) contributes anything to achievement. In other words, even if Yamamoto had taken the same High, Middle and Low IQ subgroups from his sample of quite high intelligence: (mean = 118.32) regardless of their creativity scores, the findings would probably have been the same. In fact, as Yamamoto himself points out, there was no significant difference in the mean creativity scores of the three IQ subgroups. Under these circumstances the question of any additional contribution of creativity to achievement simply does not arise. Thus, it may equally well be argued that just as there is a threshold of intelligence beyond which increase in intelligence does not necessarily lead to increase in achievement, there is also a threshold of creativity beyond which additional creativity makes no difference to achievement. What is more, in this particular study the trend of creativity scores is such that inspite of the finding of no significant difference between the creativity scores of the three IQ subgroups, in a regression equation for predicting achievement from intelligence and creativity, there is likely to be a negative weighting for creativity. For example, in comparison with the Low IQ group, the Middle IQ group have a <u>lower</u> creativity score, yet in terms of achievement scores the Middle IQ group are significantly better than the Low IQ group. Similarly although the High IQ group also have a higher creativity score in comparison with the Middle IQ group, in terms of achievement scores the two groups are indistinguishable.

The difficulties into which studies of the threshold hypothesis have run in trying to verify it by the high-creativity versus the high IQ group comparison are further highlighted by McNemar (1964) in his criticism of Getzels and Jackson's (1962) claim that creativity contributes as much to school achievement as does IQ. It will be recalled that Getzels and Jackson had compared two groups - one being the top 20% on creativity and not on intelligence the other being in the top 20% on IQ but not on creativity. With these two groups Getzels and Jackson have shown that there was no significant difference in their school achievement inspite of a 23-point IQ difference in favour of the IQ group. McNemar quite legitimately suggests that from Getzels and Jackson's data it may just as well be argued that the additional creativity of the high creativity group makes no additional contribution to their achievement.

In his study of the threshold hypothesis, Cicirelli got round the problem raised by McNemar's criticism by modifiying, in an ingenious way, the original hypothesis as proposed by J.E. Anderson (1960). According to Cicirelli, some such modification was needed "in order to explain the situation of equivalent achievement between high IQ and high-creative groups, ... for the model would predict differential rather than equivalent achievement. Such difficulties in explanation could be eliminated by postulating that the IQ level at which creativity begins to affect achievement need not be the same IQ level beyond which further IQ has no effect on achievement" (p.304). Cicirelli suggested that it would clarify the issue better if we considered two kinds of IQ thresholds:

a maximum IQ threshold beyond which "additional IQ will not distinguish individuals in terms of their academic achievement", and a minimum IQ threshold at which "creativity will begin to distinguish individuals in terms of their academic achievement."

Cicirelli tested this hypothesis by using an elaborate factorial design in which achievement in arithmetic, language and reading was considered as the dependent variable at eight levels of IQ (from IQ 70 to 140+) and three levels of creativity (high, middle and low). California Test of Mental Maturity was used as a measure of IQ and standardised tests as measures of achievement. Creativity was assessed from Minnesota Tests of Creative Thinking (verbal and non-verbal) and scored for fluency, flexibility, originality and elaboration. On the basis of factor-analysis the scores on different creativity tests were reduced to four measures as follows:

(a) Verbal fluency-flexibility-originality-VFFO

(b) Verbal elaboration - VE

(c) Non-verbal fluency-flexibility-originality - NVFFO

(d) Non-verbal elaboration - NE

On the basis of this study, Cicirelli concludes: "the hypotheses of interaction and thresholds were not supported, although there was some evidence for a maximum IQ threshold at IQ level 130-139 in the case of language achievement. While the relationship of creativity and achievement was a weak one, the form of the relationship was such that IQ and creativity were additive and linear in their effect on academic achievements" (p.308).

It may also be pointed out that correlation between IQ and the four creativity scores ranged from .09 to .24 (Average = .192), all significant at the five percent level. This is not surprising in view of the large sample size (N = 609) involved. The mean IQ of the group was 112.2, standard deviation 14.4. Thus, in regard to the question of IQ-DT relationship at this level of IQ a significant (though low) correlation between the two variables is in the expected direction. That is, the IQ level of this group is below the threshold at which such correlations approach the zero point. However, there seems to be no evidence in this study for a minimum or maximum IQ threshold to support the hypothesis of an independent contribution of creativity to achievement. As Cicirelli suggests, perhaps it is other factors like family background, cultural environment and teaching methods which may be affecting both achievement and creativity.

A study by Edwards and Tyler (1965) reports similar negative findings regarding the contribution of creativity to achievement even beyond the hypothesised IQ threshold of 120. These authors compared the school achievement of a "High SCAT" group (pupils scoring in the top third of the School and College Achievement Test, but not on the two creativity tests) with that of a "Twice Talented" Group (those in the top third of SCAT <u>and</u> Creativity), both groups having an estimated IQ of 123*.

---

*IQ was estimated from SCAT scores on the basis of previous knowledge from another study that the mean IQ of the pupils in that area was 100 with a standard deviation of 15. A correlation of .63 between SCAT and Raven's Progression Matrices is also reported for the group studied. Thus even allowing for errors in estimating the level of IQ seems high enough to justify comparison between groups for testing the threshold hypothesis.

Measures of school achievement were a standardised test (Sequential Tests of Educational Progress - STEP) and school Grade Point Average (GPA). Analysis of variance showed that on STEP there was no siginificant difference between the two groups but on GPA the "Twice Talented" group were significantly lower than the "High SCAT" group. Edwards and Tyler refer to Torrance's idea, that, creative pupils have difficulties with school requirements, curricula and interpersonal relationships, as a possible explanation for the lower GPA of the "Twice Talented" group. It may also be that the lower GPA of this group reflects teachers' prejudice against it, in so far as GPA is based on teachers' assessment whereas STEP, which is a standardised objective test shows no difference between the two groups as the subjective element of prejudice does not enter in it.

Some indirect evidence for the threshold hypothesis comes from a study by Clark, Veldman and Thorpe (1965) of "Convergent and Divergent Thinking Abilities of Talented Adolescents." On the California Test of Mental Maturity the mean IQ of 189 junior high school pupils was 125 with a standard deviation of 12. Divergent thinking tests were Guilford's Consequences, Common Situations and Seeing Problems, scored for originality. The correlation between IQ and divergent thinking at this level was .04. Thus, we have evidence that for this group of high intelligence pupils IQ and divergent thinking are "essentially independent". Evidence regarding the contribution of divergent thinking to achievement is also clear. Clark et al divided the IQ and divergent thinking scores of the group at the median to obtain a two-way classification for the analysis of variance on certain achievement and personality variable

For the two measures of verbal facility, a word fluency test and an objective, standardised reading test, the authors report finding "a highly significant ($p < .001$) mean effect for the two levels of divergent thinking ability. Subjects classified as high an the divergent thinking variable had significantly higher word fluency and reading scores than subjects classified as low" (p.160).

In Britain, Haddon and Lytton (1968) studied the threshold hypothesis with special reference to the effect of school atmosphere on divergent thinking ability. The subjects were primary school children in two "formal" and two "informal" schools with a mean VRQ of 101.75 and 101.14 respectively. Divergent thinking tests used were verbal and non-verbal tests from the Torrance tests and a new Block Printing test devised by the authors. The two hypotheses under consideration were:

"(a) Mean scores on divergent tests would be significantly higher in informal schools than in formal schools.

(b) Correlation between VRQ and divergent thinking abilities would decrease as the mean VRQ and mean divergent tests scores of subgroups rose, but the values obtained would throughout be higher in informal schools" (p.172).

Haddon and Lytton used three cut-off points on the VRQ variable - below 100, above 100 and above 115, but oddly enough, in the "above 100" category they also included the "above 115" group so that the correlations reported at different levels are not all for independent groups. With this qualification, the findings are in the expected direction, i.e. the value of correlation decreases as we go up the VRQ scale.

The same pattern of correlation emerges when a high-low dichotomy is applied to the divergent thinking score and correlations are computed with VRQ for these two groups. For the low divergent score group the correlation with VRQ is higher (r = .440) than for the high group (r = .230).

One noteworthy point about this study is that even beyond a low cut-off point of VRQ 100 the correlation between divergent thinking and VRQ drops from.512 (for the "below 100" group) to .164. However, when we look at the value of correlations for the formal and informal groups separately, it is clear that the low correlation of the "above 100" group in the formal schools (r = .059) is pulling down the correlation at this level when all the schools are combined. Thus, even at the same VRQ level there are considerable inter-school differences in the correlational analysis and the trend of correlations confirms Haddon and Lytton's "subsidiary hypothesis ... that the correlation would be somewhat higher in the informal schools".

These inter-school differences and the fact that Haddon and Lytton do not consider the effect of restriction of range on the correlations, renders their findings rather inconclusive. For example when they give VRQ/DT correlations for the formal and informal groups separately the number of subjects at the 115+ VRQ level is rather small (formal schools, n = 22; informal schools n = 13) and it is possible that the low correlations are purely a statistical artifact.

This problem of the effects of restriction of range on correlations is very crucial in any investigation of the threshold hypothesi

Although, it is possible that for certain psychological or sociological reasons the statistical prediction of low correlations with restriction of range will not be obtained. For example, Hasan and Butcher (1966) in their replication of Getzels and Jackson's study found that although the group they studied had a smaller standard deviation on VRQ than Getzels and Jackson's group, the correlation between VRQ and DT in the Scottish sample was much higher than in the American one. A psychological explanation of this finding was offered in terms of the more representative nature of the Scottish sample. Swift (1967) and Wiseman (1967) have also reported similar increases in the value of r when their samples were divided into smaller subgroups, thereby restricting the range of such variables as social class. Swift explains this in terms of the effects of moderator variables such as different parental value-systems and aspirations within one social class. He suggests that "it is better to treat social classes as social contexts in which particular factors may have special effects upon cognition and motivation" (p.10). In a replication of Haddon and Lytton's (1968) study with secondary school children, Lytton and Cotton (1969) did not find any evidence for the threshold hypothesis. They used the same tests and research design as in the earlier study with one exception: the mean VRQ of the secondary group was higher (Formal schools 111.13, Informal 112.05) than of the primary group, which was close to 100 in both types of school. Under these circumstances the threshold effect should have shown better, but it did not. The authors report that "far from increasing with the lower VRQ group, the correlation for this group practically vanished and, in actual fact, became negative" (p.189).

More recently, Bennett (1973) has also reported negative findings regarding the threshold hypothesis. Neither of these studies consider the effect of restriction of range on the correlations at different levels.

Two other studies regarding the IQ-DT relationship with reference to the threshold hypothesis must be mentioned as they have dealt with the restriction of range problem by employing various statistical devices. Ginsburg and Whittemore (1968) attempted to overcome this difficulty by not using the correlation coefficient as a measure of association between their intelligence and divergent thinking variables. Instead, they divided up their sample at different points on the intelligence variable, keeping about equal numbers in the subgroups and then obtained the Average Squared Deviations* of the Remote Associates Test (RAT) as the dependent variable, for each level of intelligence. The point at issue was whether the variance of RAT scores showed a marked increase at the higher levels of intelligence thereby demonstrating that "measures of creativity begin to differ from measures of IQ only above a certain level of intelligence" (p.133). Moving from the lowest to the highest level of intelligence, these RAT Average Squared Deviations were:

Low ← 20.33, 19.64, 21.65, 21.16, 15.60, 25.09 → High

---

*Computed as $Y\text{-}Y'^{2}/n$, where Y = RAT score i.e. the dependent creativity measure, Y' = the mean RAT score of the group within a particular band or level of intelligence and n the number of individuals within the intelligence band.

On the basis of these figures Ginsburg and Whittemore concluded that "the data presented do not support the hypothesis that RAT scores show more scatter in the higher segments of the verbal intelligence range than in the rest of that range" (p.135).

In view of Mednick and Mednick's (1965) finding of the better predictive validity of RAT rather than intelligence measures for creative performance, Ginsburg and Whittemore consider that their own results pose "an interesting problem in regard to RAT as a measure of creative ability". They come to the equivocal conclusion that "despite the persistent finding of a moderate, positive and reasonably homoscedastic relationship between RAT and verbal intelligence measures, the two appear to be measuring different (albeit overlapping) abilities" (p.136). Jackson and Messick have pointed out that RAT can hardly be considered an open-ended test since it has one predermined right answer. Cropley (1966), Hargreaves and Bolton (1972), Ketcham and Khieralla (1962) have reported finding significant positive correlations between RAT and intelligence measures. Ginsburg and Whittemore also quote a correlation of .60 between RAT and an ACER intelligence test computed from data gathered in Australia. It looks as though inspite of their use of a statistical procedure which got round the restriction of range problem, Ginsburg and Whittemore may have failed to find evidence to support the threshold hypothesis due to the unsuitability of the divergent thinking measure they chose.

In a study specifically designed to find out the effects of restriction of range and test unreliability on the IQ-DT correlation, Yamamoto (1965) used statistical formulae to correct the obtained subgroup correlations. He divided up his sample at four levels of IQ (below 90, 90-110, 110-130 and 130+) and gives the means and standard deviations for the whole sample as well as for the subgroups, so that we do get some idea of the extent to which restriction of range occurs. But, when it come to looking at the correlation between IQ and divergent thinking at the different IQ levels, Yamamoto carries out this analysis on a subsample of nineteen subjects "randomly selected" from each of the above mentioned subgroups. It is impossible to say to what extent the range was further restricted for this yet smaller group of subjects as the only statistics reported for them are the uncorrected and corrected correlations with their respective confidence intervals. In the absence of the actual values of standard deviations for the smaller subsample, Yamamoto uses the standard deviations of the initial subgroups to obtain an estimate of the amount of restriction that occurred. From this analysis Yamamoto concludes that "there was a consistent decrease in the size of correlation as the IQ level of subgroups became higher" (p.304).

The method used by Yamamoto to deal with the restriction of range problem raises as many new issues as it solves. For example, blowing up correlations by statistical methods also blows up whatever errors of measurement there are in the original data. To do this on a subsample for which means and standard deviations are unknown, increases the margin of error further and not much confidence can be put in the final value of correlations obtained in this way.

CHAPTER III

THE PREDICTIVE VALIDITY OF DIVERGENT THINKING TESTS

In view of the inconclusiveness of studies attempting to establish the convergent and discriminant validity of divergent thinking tests, and of those attempting to verify the threshold hypothesis, another line of research has been to look for the convergent-divergent distinction in the comparative predictive efficiency of the two types of tests, purporting to represent these modes of thinking. The reasoning behind this is that if divergent thinking tests can be shown to make a contribution to certain criteria of achievement, over and above the contribution made by IQ, then divergent thinking may be considered a cognitive style in its own right. Clearly, this approach is related to an aspect of the threshold hypothesis discussed earlier, and a few studies (Cicirelli 1965, Edwards and Tyler 1965, Yamamoto 1964b) have already been mentioned in the previous chapter. There are, however, other studies which have looked at this question outside the framework of the threshold hypothesis. These studies will be considered here to see what evidence there is for the unique contribution of divergent thinking ability to actual achievement.

It should be pointed out that there are two different kinds of achievements under discussion in these studies. Most of them have used academic achievement in school or college derived either from teachers' assessment or from standardised achievement tests, as the criterion of achievement. Studies of this kind will be considered fi

The other set of studies have come from authors who have argued that the above definition of achievement is too narrow and artificial. According to this view, it is suggested that achievements outside the purely academic field, or "non-academic accomplishments", as they are often referred to, ought also to be considered as a valid criterion of achievement. It is further suggested that scores based on divergent thinking tests may be better predictors of such a criterion than conventional IQ is. Studies of this kind will be considered in detail later.

Divergent Thinking and Academic Attainment

One of the main conclusions in Getzels and Jackson's (1962) study of the "highly intelligent" and the "highly creative" students was that although these two groups differed significantly in terms of IQ, there was no difference between them when their school achievement was compared. Not only were the "highly creative" group equal in achievement to the "highly intelligent" group who had a significantly higher IQ, but, what is more noteworthy is the fact that they were significantly better in achievement when compared with the rest of the sample from whom the "highly creatives" did not differ in IQ. From this finding Getzels and Jackson argued that "the cognitive functions assessed by our creativity battery accounted for a significant portion of the variance in school achievement."

Apart from the general controversy regarding their selection of experimental groups, tests and subjects, which is thoroughly documented elsewhere (Burt 1962, de Mille and Merrifield 1962, Heim 1970, Marsh 1964, Thorndike 1962, Vernon 1964 and Wallach 1970) McNemar's/

McNemar's specific criticism regarding the failure of Getzels and Jackson's study to provide evidence for a predicted differential performance (i.e. better achievement of the "highly creative" group) of the two groups has been mentioned in the previous chapter. But, McNemar's (1964) criticism itself fails to take into account the fact that the "highly creative" group did perform significantly better than the rest of the sample from whom they did not differ in IQ, so it may be argued that it was their additional creativity which contributed to their superior achievement. In other words, we may have in Getzels and Jackson's data, based on a group highly selected for intelligence, evidence for Cicirelli's (1965) minimum IQ threshold, at which "creativity will begin to distinguish individuals in terms of their academic achievement". Also, as Flescher (1963) has pointed out, it is doubtful whether differential achievement should have been expected from Getzels and Jackson's "highly creative" group since standardised objective achievement tests were used for obtaining achievement scores. On such tests the highly creative group could hardly have had an opportunity to show the kinds of ability they were best at i.e. imaginative and original self-expression.

It is true, as McNemar points out that Getzels and Jackson omit an important piece of information by not giving the creativity score also of the three groups (the two experimental groups and the rest of the sample) as this information would have helped in deciding more conclusively whether creativity does make a contribution to achievement. For example on the basis of the way in which the two experimental groups are formed it would be expected that the "highly creative" group would have a significantly higher creativity score than the rest of the sample and the "highly intelligent" group.

In such a situation, the superior achievement of the "highly creative" group vis-a-vis the rest of the sample could legitimately be attributed to their greater creativity.

Another omission for which Getzels and Jackson have been criticised is that of the "high-high" group. In a partial replication of the American study with Scottish children, Hasan and Butcher (1966) included this group also in their analysis. Their sample was more representative in terms of intelligence (mean = 102, sd = 12) and more homogeneous in age in comparison with Getzels and Jackson's group. With these differences between the two studies, Hasan and Butcher found that their "high creativity" group obtained significantly lower scores than the high IQ group on measures of English and Arithmetic attainment. As had been predicted by critics of Getzels and Jackson for omitting the high-high group, in the Scottish study this group did significantly better in English than either of the groups which was high on one variable only, but in Arithmetic attainment the high-highs were no better than the high VRQs. Considering the highly verbal bias of the divergent thinking tests used by Hasan and Butcher, it is not surprising that a creativity aggregate based on these tests contributed to attainment in English more than it did in Arithmetic. The correlation between creativity aggregate and English and Arithmetic attainment also confirm this relationship (r = .758 and .615 respectively). Thus, the Scottish study failed to replicate the findings of Getzels and Jackson regarding the contribution of creativity to school attainment.

In another replication with four experimental groups (High-IQ, High-Creativity, High IQ and High Creativity, Low IQ and Low Creativity) and using general anxiety as a moderator variable, Flescher (1963) also found no evidence for the independent contribution of creativity to achievement. Instead, he found that IQ and achievement tests were highly correlated ($r = .79$) and in the analysis of variance with achievement as the dependent variable the main effect of intelligence was significant beyond .01 while the F ratios for creativity or for interaction of creativity and intelligence were non-significant. A similar analysis with anxiety showed that neither IQ nor creativity had a significant effect on this variable either independently or jointly. Therefore the possibility of the effect of anxiety as an intervening variable was ruled out. From these findings Flescher concluded that "as anticipated, intelligence is portrayed as a powerful determinant of academic achievement" (p.257).

There are two features of Flescher's study which ought to be mentioned here as they underline the significance of his findings. Firstly, the correlation between IQ and divergent thinking total was only of the order of +.04. This is not unexpected since he does make it clear that the mean IQ of his subject was "in the upper segments" of the scale. Perhaps it was at the level at which convergent and divergent abilities are no longer related. This is further confirmed by the fact that even his two low IQ groups (i.e. High-Creativity and Low Creativity-Low IQ) had a mean IQ of 115 and the High-High group went up to a mean of 142.

But the second important aspect of Flescher's study is that even at this high level of IQ the effects of creativity on achievement failed to show inspite of the prediction of the threshold hypothesis that at higher levels of IQ creativity is more likely to make a contribution to achievement. Could it be that in earlier studies like Getzels and Jackson's, and Torrance's (1962) and Yamamoto's (1961) replication of them a significant effect of creativity on achievement was found precisely because IQ and creativity were themselves positively related, whereas they are not in this study? On theoretical ground McNemar (1965) has argued for this view and Flescher's findings bear it out.

In his study of the reliability and validity of creativity tests, Wodtke (1964) also came to a negative conclusion regarding both these issues. As criteria of achievement he used an imagination story (a divergent task) and the Luchins' Water Jar Test (WJT, a convergent task). Lorge-Thorndike IQ and Torrance's verbal and non-verbal creativity tests were used as predictors. The aim of the study was to see if either of the criteria could be predicted by the creativity tests holding IQ constant by a partial correlation technique. None of the partial correlations for predicting the WJT from the creativity tests was significant whereas IQ was found to be positively correlated with this criterion. With the imaginative story the findings were more mixed; in grade four IQ predicted this criterion also better than did the creativity tests, but in grade five IQ and creativity predicted the imaginative story score equally well. Wodtke considers that these negative findings can be explained partly in terms of the generally low reliability of the creativity predictors as well as of the criterion measures, but even after correcting the correlations for attenuation due to unreliability, the pattern of relationships was still unchanged.

Mixed findings rather similar to Wodtke's regarding the predictive validity of divergent thinking tests for writing imaginative stories have recently been reported by Bennett (1973). He used rescaled teachers' rank-orders in English and assessment of imaginative stories for fluency (in terms of total word count) and for over-all impression marking, as the criteria of achievement. Predictors in this study were a verbal reasoning test and five divergent thinking tests. A correlational analysis of these variables showed that the verbal reasoning test and rescaled teachers' marks in English were highly correlated (.90 and .87 for boys and girls respectively), whereas the aggregate of divergent thinking tests had much lower correlations with this criterion (.57 and .55 for boys and girls respectively). In predicting the more creative criterion, i.e. the imaginative story, verbal reasoning scores and divergent thinking scores show about equal efficiency. Correlations between verbal reasoning and the story range from .41 to .49 and those between the divergent thinking aggregate and the story from .43 to .55.

There is a marked drop in the above correlations between VRQ and imaginative story in comparison with the correlations obtained between VRQ and rescaled teachers' estimates of English attainment. It cannot be accounted for in terms of the low reliability of the criterion as Bennett reports that "mark/re-mark and inter-marker correlations both exceed .70" for the imaginative story. It is more likely to be due to a lack of congruence in the subject-matter or content of the predictor and criterion variables.

The verbal reasoning test is an objective one-right-answer type test emphasising logical thinking and reasoning whereas the story was explicitly assessed for the use of imagination and good ideas, poor spelling and grammar were not penalized. Considering these differences, the reported correlations between VRQ and the story (ranging from .41 to .55) are not surprising. Perhaps it is the exceptionally high correlations between VRQ and teachers' rescaled estimate of English attainment which require some explanation. Bennett suggests that since teachers' rank orders in English were scaled against the verbal reasoning test this may have artificially increased the obtained correlations. But this does not follow from the scaling procedure used (Yates and Pidgeon 1957) unless there is already a close correspondence between teachers' estimates and VRQ. A more plausible explanation seems to be that a "halo" effect is operating in the teachers' estimates. This could happen if the teachers knew the pupils' VRQ from school records, in which case the rank order they assigned for English attainment may have been influenced by this information. If this was not the case then the correlations are very impressive indeed.

In a factor-analysis the imaginative story was isolated as factor three on which three divergent tests of associational fluency had "low to moderate loadings" (.30, .47 and .28 respectively). At the same time the impression marks for the story had also been found to load .31 on the first factor of convergent ability. Thus, it is clear that in terms of comparative predictive efficiency even for such an open-ended criterion as the imaginative story, divergent abilities show no clear superiority although when considered alongside the verbal reasoning test they do predict this criterion slightly better than they predict conventional exam marks in English.

In their follow-up of the pupils they had tested at the primary level to study the effects on divergent thinking, of formal and informal school climate, Haddon and Lytton (1971) also included the question of the predictive validity of divergent thinking tests over a four year span. As in most other studies, they used standardized tests of English and Arithmetic-Mathematics given at age 15 as measures of achievement. Verbal Reasoning Quotients(VRQ) and divergent thinking scores from the earlier study when the subjects were eleven years old were the predictors. For reporting correlations, Modern and Comprehensive schools are grouped together and Grammar schools are separate. This was done in view of the differences in the magnitude of correlations for these two sets of schools. The findings were clearly against any significant contribution of divergent thinking to attainment when compared with the contribution made by VRQ. For example, the correlations between VRQ and English/Mathematics for both types of school ranged from .29 to .83 and those between divergent thinking and English/Mathematics were from .01 to .38.

On the whole all correlations for the Grammar school were lower than for the Modern and Comprehensives together. Haddon and Lytton point out that to some extent this may be due to the highly selected nature of the Grammar school population and the consequent restriction of range on both the predictor and criterion variables, but they also suggest the "additional possibility ... that the effect of four years schooling is greater in the Grammar schools, and hence attainment there can be less easily predicted by the measurements of ability variables, whether convergent or divergent" (p.145).

It is not clear how four years of Grammar school education could make predictability of performance from tests more difficult other than for reasons of greater homogeniety of the group and therefore the lack of sufficient discriminating power in tests normed for the total population. Whatever the explanation for generally lower correlations in the Grammar schools, the trend of considerably higher correlations between VRQ and attainment than between divergent thinking and attainment is clearly present in both sets of schools.

From the studies considered so far, the evidence for the predictive validity of divergent thinking tests has been largely negative. But, a number of studies have been reported which have found more positive evidence for Getzels and Jackson's view that scores on divergent thinking tests make a significant contribution to the variance of school achievement measures. The earliest of these was a series of replications carried out by Torrance (1960) at eight different schools. Six of these confirmed Getzels and Jackson's findings and for the two which did not Torrance invoked the threshold hypothesis to explain the discrepant findings; another explanation given was that the schools were "known for their emphasis on traditional virtues in education". Following Getzels and Jackson's procedure, Torrance had formed his highly intelligent and highly creative experimental groups by taking those who were amongst the top 20% on the one variable but not on the other. Those who were in the top 20% on both measures were exluded from the study. Thus the comparison was between two extreme groups and Torrance reports that inspite of an average difference of twenty-five IQ points between the two groups, no statistically significant differences were found in six of the eight schools when school attainment of the groups on objective achievement tests was compared.

As Wallach has pointed out the tests of divergent thinking used by Torrance and his colleagues have repeatedly been found to have a significant positive relationship with intelligence and therefore the contribution of intelligence to achievement even in this extreme group comparison cannot be ruled out.

Yamamoto has also reported a number of studies regarding the role of divergent thinking in educational achievement (Yamamoto, 1964a, 1964b, and 1964c). In the first of these (1964a) he included a high-high group also, besides the two Getzels and Jackson type high-IQ and high-creativity groups. Using an analysis of variance procedure he found no significant difference in the achievement scores of the three groups and considered Getzels and Jackson's finding regarding the contribution of creativity to achievement confirmed. But, he did point out that the whole sample from which the experimental groups were taken were already highly selected for intelligence (mean IQ 118, sd 15) and "this might have had some effect on their high performance on achievement tests" (p.788). What Yamamoto does not comment on is the fact that even the high creativity/low-IQ group had a mean IQ of 122 (sd 8.5). This is above the postulated threshold of 120 at which the differential effect of creativity on performance (i.e. higher achievement) ought to show, but it does not. That creativity was not contributing to achievement even at this high level of IQ was further confirmed in a reanalysis of the same data (Yamamoto 1965a) in which intelligence was controlled for by an analysis of covariance and it was found that there was still no significant difference in the achievement scores of the high-IQ, high-creativity and high-high groups. Again, the high-creativity group did no better than the other two groups.

Cline, Richards and Abe (1962) and Cline, Richards and Needham (1963) have investigated the predictive validity of divergent thinking tests using all subjects in their sample instead of extreme groups, the sample being more representative of the general population than was the case in the studies discussed earlier (Getzels and Jackson 1962, Torrance 1960, Yamamoto 1964a). In both the studies by Cline et al the mean IQ for males and females on the California Mental Maturity Invetory, ranged from 98.67 to 101.22 and the sd from 11.64 to 12.32. Thus in both cases it was a fairly average group of high school seniors on whom the studies were based. Five of Guilford's divergent thinking tests (namely, Consequences, Word Association, Hidden Figures, Brick Uses and Match Problems) were also used as predictors alongside IQ. In the earlier study by Cline et al (1962) the criterion was academic performance in high school in terms of grade point average over a period of three years. Beta weights of each divergent thinking test and IQ were computed for the GPA criterion. Multiple correlations are also reported between the divergent battery alone and GPA and between divergent battery plus IQ and GPA. Multiple correlations obtained in this way are as follows:

	Boys		Girls	
	DT only	DT + IQ	DT only	DT + IQ
Multiple r with GPA	.65	.69	.63	.68

In view of the small increase in the multiple correlation when IQ is added to the divergent thinking tests, Cline et al (1962) conclude that "the criterion variance accounted for by creativity tests is to a substantial degree independent of the variance accounted for by the IQ test" (p.784).

However, looking at the beta weights and the intercorrelations among the predictor and criterion variables it is clear that there are other possible explanations for the small increase in the multiple correlation when IQ is added to divergent thinking tests. For example, Wallach (1970) has shown that in this study predictor variables are themselves positively correlated (average r between IQ and DT for boys is .35, for girls .32), and therefore the multiple correlation between divergent tests and GPA is not an index of the independent contribution of divergent thinking to school achievement. In fact, when we compare the correlations between IQ and GPA and DT and GPA we find them to be as follows:

	Boys	Girls
IQ - GPA correlations	.57	.55
DT - GPA correlations	from .13 to .37	from .14 to .36

On the basis of these figures the superiority of IQ over DT as a predictor of GPA is clear. Also, it is not legitimate to compare the multiple correlations obtained with and without IQ when there is only one measure of IQ being used while there are five different divergent thinking tests. And, if we look at the individual beta weights for the divergent thinking tests and IQ we find that the beta weights for IQ are highest of all the predictors. They are + .3220 and + .3397 for boys and girls respectively; whereas the beta weights for the divergent thinking tests range from - .0892 to + .2359 for boys and from - .0082 to + .2548 for girls. This was only to be expected from the first order correlations which are higher between IQ and GPA than they are between any individual divergent thinking test and GPA for both the sexes. In view of these comparative figures overlooked by Cline et al their conclusions regarding the predictive validity of divergent thinking tests hardly seem warranted.

In the second study by Cline et al (1963) basically the same research design and modes of analysis were used with the exception that it was achievement in science only that was chosen as the criterion. But in this study, measures other than the GPA in science courses were also used to obtain five criteria of science achievement. Four of the other criterion variables were : STEP Science Test, Teacher Rating of Science Potential, Number of Science Courses Taken and Involvement with Science. In view of the complexity of these criteria, findings are more mixed in this study but the sex difference in the pattern of beta weights reported in the earlier study becomes even clearer in this one. Addition of IQ to divergent thinking tests makes very little difference to the multiple correlations for boys for any of the five criteria. In fact there are two negative beta weight for IQ but they are not high enough to actually reduce the multiple r. For girls the beta weights for IQ are highest in every case when compared with the divergent thinking tests and this is reflected in considerably higher multiple correlations with each of the criterion variables when IQ is included as a predictor. Again, Wallach (1970) reports average correlations of .35 and .33 between IQ and divergent thinking for boys and girls respectively and the comments made about the effects of this on the findings of the earlier study (Cline et al 1962) also apply here. Additionally, regarding the finding that boys are penalized by science teachers (mostly male) in grading and in rating, in comparison with girls, it may be due to the higher expectations of performance in science from boys so that what is considered to be good performance for a girl may only be average for a boy.

Perhaps these differential expectations are related to socio-cultural conceptions of different activities and areas of achievement being more or less appropriate for the two sexes.

The conclusion to be drawn from the studies reviewed in this section is clearly negative. In no case has the predictive validity of divergent thinking tests for academic attainment been demonstrated unequivocally. To some extent this is not surprising since the measures of achievement have been fairly conventional types of academic attainments, based on a prescribed curriculum with clearly defined correct answers or solutions to the learning tasks. Often achievement scores for these studies have been obtained from multiple-choice type objective tests. Even when they are not of the type where a predetermined right answer is expected from pupils, they are based on school or college examination marks which are strongly biased towards the correct reproduction of material taught during lessons/lectures or learnt from books. It has therefore been suggested that the predictive validity of divergent thinking tests, which require the ability to produce ideas rather than reproduce them, can hardly be expected to be high if tasks of a convergent type are used as criteria. But we have seen that even in studies where an open-ended task like writing an imaginative story or autobiography was the criterion (Bennett 1973, Flescher 1963, Wodtke 1964), divergent thinking tests did not show any superiority over IQ measures as predictors.

It may be that these negative findings are largely due to the fact that even the studies which have used fairly open-ended criterion measures have been conducted under conditions of external constraints so far as the subjects have been concerned.

That is to say, subjects have been asked to perform certain tasks, like writing a story or autobiography at a certain time chosen by the researcher or the school authorities and the question of motivation or interest on the part of the subject has been generally overlooked. In a recent study considered to be a validation of the creativity-intelligence distinction, Wallach and Wing (1969) have argued that by the very nature of abilities involved in doing divergent thinking tests, they are better predictors of tasks and activities undertaken out of one's own interest and initiative rather than those imposed externally as a requirement. Thus, Wallach and Wing extended their definition of achievement to include "non-academic accomplishments" in the field of art, music, science, literature etc. The rationale and details of studies of the kind illustrated by Wallach and Wing's work, will be discussed in the section which follows, to see if evidence for the predictive validity of divergent thinking tests can be found in activities and achievements generally considered to be outside the traditional academic sphere.

Divergent Thinking and Non-Academic Accomplishments

In the literature dealing with the question of non-academic accomplishments, this term is used to indicate achievement in fields of activity not included in the prescribed, examinable curriculum of schools and colleges. It refers to what have been known to teachers as "extra-curricular activities" with one important exception, namely, achievement in the field of games, sports, athletics and other aspects of physical education are usually excluded.

Included in the term are pursuits like creative writing, art and craft work, debating and acting, photography, making mechanical/electrical models, nature study in the form of bird-watching or collecting specimens etc., as well as social service and leadership in school, college or community. For an activity to be included in this category a necessary condition is that it should have been pursued on one's own initiative and interest, although no doubt other persons may have helped and advised from time to time. From the types of activities mentioned above, it is clear that any of them can be pursued in an academic manner, if by academic we mean a systematic and informed way of doing things. But in this context "non-academic" is used to emphasise the fact that these pursuits are outside the formal requirements of educational institutions. It is further assumed that opportunities for such pursuits are available in a greater or lesser degree to most students.

There are two lines of reasoning involved in the argumentfor including non-academic accomplishments as a criterion of achievement alongside the more traditional academic attainment criterion, especially if tests other than the conventional intelligence tests are being used as predictors. Firstly, there is the question of the theoretical and empirical validity of the criterion itself. For example, in an extensive review of literature concerning the relationship of college grades to level of success in adult life in the field of business, teaching, engineering, scientific research, medicine and overall eminence, Hoyt (1965) concluded that "college grades have no more than a very modest correlation with adult success, no matter how defined. Refinements in experimental methodology are extremely unlikely to alter that generalisation; at best they may determine some of the conditions under which a low positive, rather than a zero, correlation is obtained" (p.45).

The implication of this statement is that the usual criterion of academic achievement in predictive validity studies is not in this case a very meaningful one.

A related argument, put forward most forcefully by Dunnette (1963) is that the practice of selecting a single criterion oversimplifies the complexity of human performance. He suggests that "applied psychologists should give more emphasis to construct validation and make an effort to learn more about the meaning of test scores and other predictors in terms of multiple dimensions of behaviour" (p.251). In commenting upon the undesirability of equating giftedness with high IQ, Getzels and Jackson (1963) also stated that "if we moved the focus of inquiry from the classroom setting, we might identify qualities defining giftedness for other situations just as the IQ did for the classroom. Indeed, without shifting our focus of inquiry if we only modified the conventional criteria of achievement, we might change the qualities defining giftedness even in the classroom. For example, if we recognised that learning involves the production of novelty as well as the memorisation of course content then measures of creativity as well as the IQ might become appropriate in defining characteristics of giftedness" (p.190). Thus, here we have another plea for broadening not only the criteria of achievement but the predictors too.

The second reason for extending the definition of achievement to cover non-academic accomplishments is implied in the last part of Getzels and Jackson's statement quoted above.

It is that the validity of any predictor depends on the criterion chosen to validate it against. It is hardly reasonable to expect a high predictive validity if the criterion chosen is conceptually unrelated to the predictor. In psychometric terminology, it is a question of the construct validity of the predictor. With specific reference to the present discussion, it is argued that divergent thinking tests involve productive rather than reproductive abilities and therefore would be related more to the creative pursuits of the non-academic type mentioned at the beginning of this section.

Wallach and Wing (1969) tried to verify empirically the possibilities suggested by Getzels and Jackson regarding the construct validity of creativity tests. In so doing, they brought together two areas of research that had been developing independently in the previous decade. J.L. Holland and his colleagues at the National Merit Scholarship Corporation in America had been developing questionnaires and rating scales for obtaining reliable and valid measures of non-academic accomplishments to be used as additional criteria of achievement. From this source Wallach and Wing derived their non-academic accomplishments questionnaire. Scores on this questionnaire were used as an extended criterion of achievement. To broaden the basis of prediction, Wallach and Wing used divergent thinking tests along with the IQ variable. These tests had been used in predictive validity studies, largely with negative results, as discussed in the previous section. Wallach and Wing reasoned that if the criterion of achievement included non-academic accomplishments divergent thinking tests would show better predictive validity than had been found with conventional attainment measures as the criterion.

In the studies of non-academic accomplishment carried out by Holland et al (Holland 1961, Holland and Astin 1962, Holland and Nichols 1964, Holland and Richards 1965, Richards, Holland and Lutz 1967) multiple predictors of this extended criterion have been used. Only the measures employed most frequently in these studies are described briefly here:

Potential Achievement Scales: Consists of check-lists of daily activities, hobbies, reading habits, school subjects, sports etc.

High School Achievement Scales: A check-list of unusual accomplishments in science, arts, leadership, drama, music and literature.

Vocational Preference Inventory: Contains 11 scales composed of different occupational titles such as Realistic, Intellectual, Conventional, Enterprising, Artistic etc.

Range of Experience Checklist: Items regarding places visited (museum, factory, casino) and events experienced (sports car race, summer camp etc.).

Intellectual Resources in the Home: Items included - encyclopaedia set, tape recorder, sculpturing tools, power tools, books etc.

Range of Competencies: Students chose activities they thought they could do well, such as dance, make jewellery, read blueprints, operate a tractor, use logarithm tables etc.

Other instruments used were Deferred Gratification Scale, Super-Ego Scale, Independence of Judgement Scale, Aspirations and Goals, as well as a Parental Attitude Research Instrument (PARI).

These predictors were used in different combinations in the studies mentioned above. Some studies also included such conventional measures as GPA or standardised achievement test batteries (Holland 1961, Holland and Richards 1965), but any direct measure of intelligence such as the Scholastic Aptitude Test has not usually been included amongst the predictors of non-academic accomplishments.

The general research design used in most of these studies is that of first order correlational analysis, often followed by calculation of beta weights of the different predictors and the value of multiple r as the predictors are added in the regression equation.

Summarising the main findings of this series of research studies, Holland (1966) reports predictive validities averaging .38 using records of activities and accomplishments in high school, for National Merit Scholarship Finalists. In a replication of these studies with a more representative student population the predictive validities average .40. Reliabilities, (usually KR 20) of the non-academic accomplishments questionnaire used as a criterion are reported to range from .65 to .84 at the high school level and from .44 to .80 at the college level. Holland et al draw a number of related conclusions from these studies:

1. "For samples of students of superior scholastic aptitude, creative performance* is generally unrelated to scholastic achievement and scholastic aptitude."

---

*"Creative performance" is used by Holland in this study as a synonym for scores on the non-academic accomplishments questionnaire.

In an earlier part of the same study (Holland 1961) Holland distinguishes the academic achiever: "good grades in high school appear to be a function of socialisation (citizenship and popularity) and perseverance, whereas creative performance is a function of conscious concern with high accomplishment, independence and originality" (p.143).

2. "If sponsors wish to find students who do outstanding things in college, then they should continue to make an effort to secure a better record of the student's involvements, achievements, self-conceptions and goals while he is in high school" (Holland and Richards 1964, p.65). In a later paper (Holland 1966) summarising the main findings of the research, Holland further elaborated this point: "we have only engineered what every layman and mother knows: To find out if a man is going to become an outstanding performer, simply add up his little performances as he moves through life" (p.48).

3. "Non-academic accomplishment can be assessed with moderate reliability and non-academic potential can be predicted with moderate success" (Richards, Holland and Lutz 1967).

4. Finally, Holland (1966) hoped that "these non-intellective materials will not become another hurdle in a highly selective admissions procedure. ... psychometrically it is better to use many psychological devices with moderate reliability and validity against several criteria than a few - or only one - instrument with high reliability and high validity against a single criterion" (p.49).

Werts (1967) has pointed to various methodological shortcomings in Holland and his group's research design which make the conclusions quoted above rather questionable. Following McNemar's (1964) line of argument, Wert's main criticism centres round the inappropriateness of correlational analysis in studies where there is a restriction of range in the predictor and criterion variables. That such a restriction can significantly alter correlations has been shown by Skager, Shultz and Klein (1965). With IQ and socio-economic status as predictors and quantity as well as quality of non-academic accomplishments as criteria, these authors first analysed results for two colleges separately and then together. In the separate analysis no significant correlation was found between IQ (SAT, V and M) and either the quality or the quantity criterion. But it was also observed that the range of SAT scores was considerably restricted. When data for the two colleges were combined the SAT score dispersion was close to the norm. Correlations computed on these combined data also showed a significant positive relationship between SAT and quality of non-academic accomplishments. With specific reference to the studies of Holland et al. Werts argues that the infrequent occurance of non-academic accomplishments even in the highly selected group of National Merit Finalists is bound to result in low correlations between GPA or ability measures and the criterion of non-academic accomplishment. This is particularly so if the criteria of accomplishment include achievements in the field of music , dramatics, art etc., since people particularly talented in these areas would most probably go to specialised institutions rather than to four year colleges of the type where most of this research has been carried out.

The issue is further complicated by the specific nature of non-academic accomplishments and the limited coverage of such accomplishments in a questionnaire or checklist. In such a situation even the multiple correlation model does not do full justice to the predictive power of high school grades (HSG) because some people may be utilising their academic skills in areas not covered by the questionnaire. In regard to this last point it should be pointed out that it is precisely this specific nature of non-academic accomplishments which Holland et al have used as an argument for using equally specific predictors rather than a more general measure of academic aptitude or attainment (Holland 1967). It seems therefore that Werts' emphasis on this point lends support to the line of reasoning taken by Holland et al instead of being a criticism of it.

The alternative procedure employed by Werts in dealing with the problems raised in his criticisms, is as follows: he plots the percentage of pupils checking different kinds of non-academic accomplishments who fall at various points on a nine -point scale of high school grade average. Analysed in this way his data show that "the percentage of achievers rose exponentially with rise in HSG for the three science and the three literary items" (p.200). In order to compare his findings with a study using correlational analysis only (Holland and Richards 1965), Werts also reports correlations between the various non-academic measures and high school grade average. These are generally as low as those obtained by Holland/

by Holland and Richards*, yet within this low degree of overall relationships there is a clear trend for a higher percentage of pupils at the upper end of the HSG scale (A+, A etc.) to report non-academic accomplishments in comparison with those at the lower end. Thus, Werts claims to have made a more legitimate comparison by using a procedure in which the percentage of people at different grade levels checking a particular accomplishment are compared with each other rather than with an "inappropriate reference group", i.e. students in general, as the correlational procedure of Holland and Richards had done.

Holland and Richards' (1967) reply to the criticism made by Werts rests mainly on the argument that although Werts has shown a relationship between high school grade average and non-academic accomplishments he over-looks the number of students with such accomplishments who would still be <u>missed out</u> if high grades alone formed the basis of selection. In other words, Holland and Richards are referring here to individuals who in medical terminology are called "false positives". But as Anastasi (1968) and Cronbach and Gleser (1965) in their discussion of the application of decision theory to psychological testing have pointed out, the question of how many people are wrongly accepted or rejected in a selection procedure is not a function of test validity alone. It depends also on the value of the selection ratio and base rate of occurrence. The former refers to the percentage or proportion of applicants/ testees who can be selected at a certain time. Obviously this decision is affected by such external constraints as the availability of resources in relation to the demand for them.

---

*In the Holland and Richards (1965) study the product-moment correlations ranged from -.05 to +.21. In Werts' (1967) study point-biserial correlations were computed and these range from +.01 to +.25.

The latter term, base rate, used usually in clinical psychology, refers to "the frequency of a given condition in the population to which a test is applied" (Anastasi 1968). It has been shown (Taylor and Russell 1939) that with extreme base rates (too frequent or too rare occurrence) the incremental validity of any test for clinical diagnosis is only marginal.

In view of these considerations, it is not surprising that selecting students on the basis of high school grade average does lead to a number of "false positives" in terms of non-academic accomplishments. The size of this group will increase as the selection ratio is made more astringent (e.g. selecting A+ students only). Indeed this is what Holland and Richards show when they reanalyse Werts' data for estimating the loss of talent. The percentage of students with non-academic accomplishments who are missed out increases as we move upwards from grade average B+, A-, A to A+. The second factor of a low base rate of non-academic accomplishments adversely affecting the predictive validity of high school grades in the studies of Holland et al is also implied in the restriction of range in the criterion mentioned by Werts.

In a recent study, Wallach and Wing (1969) accept that the Werts (1967) study does show "some degree of relationship between academic and non-academic accomplishments" (p.9), but they suggest that it may be an artifact of the use of grade average rather than IQ as a predictor. They proceed by arguing that there is a common element of motivation in all kinds of achievement whether academic or non-academic, and this may explain why Werts did find a relationship between these two kinds of achievement:

"Given an achievement motivation component in a person's level of academic accomplishment, we might expect a modicum of relationships with non-academic forms of accomplishment to exist since the latter may also be influenced to some minor extent by the degree of achievement striving that characterises the individual" (p.10). IQ on the other hand, comes "closer to the level of sheer intellective ability ...thus reducing - although not necessarily eliminating - the extent to which motivation for achievement as such may bear upon the results" (p.9). In their own study therefore, they use IQ instead of grade average as a predictor variable. This procedure assumes the relative independence of IQ from motivational variables and an unquestionable link between achievement and motivation. Neither of these is justifiable. For example, in his review of studies of the role of achievement motivation and academic achievement, Lavin (1965) concluded: "In general, the research does not indicate that achievement motivation is strikingly related to academic performance" (p.78). Direct studies of the relationshp between IQ and motivation have not often been carried out. The question has been studied as part of the academic achievement/motivation issue to determine the relative contribution of intelligence and motivation to achievement. In this context, McClelland (1953) has reported significant positive correlations between intelligence (ACE or SAT scores) and n Achievement. Using her method of Story Sequence Analysis Arnold (1962) has reported even higher correlations ranging between .4 and .5.

On the methodological side Wallach and Wing counter Werts' criticism regarding the inappropriateness of correlational analysis in the studies of Holland et al by comparing "groups that are higher and lower regarding intelligence for their mean relative frequencies of one or another category of non-academic accomplishments" (p.10).

They extend the scope of their study by hypothesizing that not only is there a lack of any significant relationship between intelligence and non-academic accomplishments but also that the latter are positively related to another mode of cognitive functioning i.e. creativity or divergent thinking. In view of Wallach and Kogan's (1965) earlier conclusion that the production of many and unique ideas characterises this mode of thinking most typically, Wallach and Wing also defined creativity in terms of Ideational Productivity and Ideational Uniqueness. Before the hypothesized relationships between idea producing ability and non-academic accomplishments could be shown, the independence of IQ and idea producing ability had also to be established. We have seen in the previous chapter that scoring for number and statistically rare responses does yield a maximum differentiation of divergent thinking from IQ. Accordingly, the data in this study showed considerable independence of these two variables.

Wallach and Wing's study was carried out on a volunteer sample of 502 Duke University entrants. The mean IQ of the group was: SAT(V) 619.12, sd = 70.12, SAT(M) 644.88, sd = 69.23. There was a significant sex difference on the Verbal and Mathematical scores, women did better on the former, men on the latter. For number and uniqueness of ideas, two of Wallach and Kogan's (1965) verbal (Uses, Similarities) and two non-verbal (Pattern Meanings, Line Meanings) tests were used. As a measure of academic achievement high school grades and freshman grades were utilised separately. And finally, non-academic accomplishments score was derived from a self-report questionnaire covering achievements in the field of Leadership, Art, Social Service, Literature, Dramatic Arts, Music and Science.

SAT scores and grades (high school and freshmen) were obtained from office records; open-ended tests and the questionnaire were sent to the subjects by post to be completed at home during the summer before they started their first semester at the University.

The required convergent and discriminant validity of divergent thinking tests was evident in the high intercorrelations among these tests (r ranged from .38 to .79) and in the low correlations between them and IQ (r ranged from -.07 to +.09). Having established this distinction, Wallach and Wing proceed to compare the academic and non-academic achievements of students scoring in the top third of the intelligence, ideational productivity and ideational uniqueness variables with those scoring in the lower third on these variables. In every case, findings are reported for both sexes separately as well as for the whole sample. Throughout the analysis, t tests are used to obtain the significance of difference between the dependent variable means of the high and low groups.

Inspite of an almost zero correlation between SAT and divergent thinking tests, the predictive validity of these two measures for academic achievement are very similar. For example, when the high school grades of the high and low groups are compared, the high groups on all three predictors (IQ, ideational productivity and ideational uniqueness) show significantly better grades than the low groups with only one exception. That is, on the ideational uniqueness variable there is no significant difference in the high school grades of women in the high and low groups, although the difference in mean scores is in the expected direction.

Wallach and Wing do not report the number of subjects who may have been in the high or low groups on more than one independent variable simultaneously. Considering the low correlation between IQ and ideational productivity/ideational uniqueness one would not expect too much overlap here. On the other hand, it may be that there is more overlap in these extreme groups than in the middle group which has been excluded from the analysis. If this is the case, then we are dealing with the grades of quite a few people who are appearing consistently in the high or low group whether selection is made in terms of IQ or ideational productivity or ideational uniqueness. Such a situation would explain why the findings for all the three variables are so similar. Wallach and Wing completely ignore this possibility and conclude that "ideational flow on the one hand and intelligence on the other are making relatively separate contribution to a students' high school grade level" (p.58). What we need to know before this conclusion can be accepted is the value of correlation between IQ and divergent thinking measures for the high and low groups separately.

When academic achievement at the end of freshman year is used for a similar analysis, the findings are rather mixed. The high IQ group again comes out with significantly higher achievement. For the high productivity group too there is a significant difference in the same direction if men and women are considered together; but in the analysis for the sexes separately, women in the high productivity group do not show significantly higher achievement scores. The high and low ideational uniqueness scores also fail to make any discrimination in terms of college achievement.

This difference in the findings for high school and freshman grades is better understood in the light of information provided by the authors about what these grades are based on. High school grades were based on the class-rank of an individual "expressed in relation to the size of his class". The ranks were converted to a weighted score to make them comparable to SAT scores. Such class ranks prepared for seniors reflect the students' academic standing over the high school years and are likely to be a more reliable estimate of achievement than the Freshman Quality Point Ratio used as the criterion of academic achievement in College. This criterion is based on grades given in the impersonal setting of a large college to freshmen who are not so well-known to their teachers and who may still be going through difficulties in the process of adjusting to the new environment. What is more, the grades are on a five point scale so that there is not much scope for differentiation in comparison with the ranking procedure used at the high school level. Therefore, the unreliability of college grades may be responsible for the difference reported above. Again, Wallach and Wing do not consider this possibility and explain their findings mainly in terms of the greater predictive validity of ideational productivity (relative to ideational uniqueness) for college grades.

For the present discussion, the most relevant part of Wallach and Wing's study is the one which deals with the comparative predictive power of IQ and divergent thinking in regard to non-academic accomplishments*.

---

*Strictly speaking, the foregoing discussion regarding academic achievement belongs in the previous section, but it seemed more appropriate to leave it out then and consider it after the general background of Wallach and Wing's study had been explained.

To get a complete picture of the main findings in this area, it is best to summarise them in tabular form. In the table which follows, M, F and T stand for males, females and total sample respectively, and ns indicates a finding of no significant difference between means of the top and lower third on IQ, productivity and uniqueness. Where a significant difference was found, the probability level is given:

Table III.1 : Summary of Wallach and Wing's (1969) findings regarding mean non-academic accomplishment scores of groups high and low on IQ, Ideational Productivity and Ideational Uniquenesss.

Non-Academic Accomplishments	IQ			Ideational Productivity			Ideational Uniqueness		
	M	F	T	M	F	T	M	F	T
Leadership	ns	ns	ns	.01	.01	.001	ns	ns	ns
Art	ns	ns	ns	.01	.02	.001	.01	.01	.001
Social Service	ns	ns	ns	ns	ns	ns	ns	ns	ns
Writing	ns	ns	ns	.01	ns	.001	.02	.001	.001
Dramatics	ns	ns	ns	ns	ns	ns	ns	ns	ns
Music	ns	ns	ns	ns	ns	ns	ns	ns	ns
Science	ns	ns	ns	.01	.02	.001	ns	.05	.01

From the above table, Wallach and Wing's first thesis regarding the lack of any significant relationship between IQ and non-academic accomplishments stands confirmed. But the second thesis regarding the predictive validity of ideational productivity and ideational uniqueness finds only partial confirmation. If we look under the productivity and uniqueness columns, the non-significant differences are just over half (23 ns, 19 $p < .05$ or better).

Clearly, in these cases, the divergent thinking predictors like the IQ, failed to discriminate between the high and low non-academic achiever. On the other hand, it is also true that some of these predictors have shown better discrimination than IQ, which had totally failed in this respect. Also, if we look at the results Wallach and Wing present for the total score on this questionnaire, it is clear that with only one exception (males in the uniqueness table) the productivity and uniqueness measures discriminate quite significantly ($p < .001$) between non-academic achievers, whereas all differences for intelligence are again non-significant. Considering that the different possible fields of accomplishment included in the questionnaire are covered by only three to six "items", it is not surprising that a number of differences in the table above turned out to be non-significant. Therefore, comparison of total scores seems a more reasonable procedure and when this is done the better predictive validity of productivity and uniqueness measures for non-academic accomplishments emerges clearly. It should also be noted that amongst the predictors, intelligence is a composite of verbal and mathematical scores (which in turn are derived from scores on a large number of items) whereas the divergent thinking predictors are based on a total of twelve verbal and non-verbal "items". In view of these limitations of the productivity and uniqueness measures, it is possible that with a longer open-ended test, Wallach and Wing's second thesis of a significant relationship. between creativity and non-academic accomplishment would have found unequivocal support.

Another weakness in Wallach and Wing's study is the possible existence of a volunteer bias in the non-academic criterion and the divergent predictors, which makes the meaning of the reported significance levels even for the total score on the questionnaire doubtful. Wallach and Wing did check that "no volunteer bias with regard to intelligence was operating" (p.30), but the significant differences reported are in the mean scores of the high and low divergent groups on the questionnaire. Since Wallach and Wing found a near zero correlation between IQ and the productivity and uniqueness measures, there is no guarantee that volunteer bias is not operating with regard to this predictor variable. We also do not know to what extent this bias is present in the criterion scores. For example, it is possible that only those who had an accomplishment to report responded and the non-respondents were on the whole a low-accomplishment group.

Two studies with secondary school students (Cropley 1972, Kogan and Pankove 1972) have appeared recently which provide some more qualified support for Wallach and Wing's thesis. It is noteworthy that both these studies are truly predictive in the sense that divergent thinking scores obtained five years before are used in the correlational analysis, although Kogan and Pankove also include concurrent creativity scores in the multiple correlation. These studies have been described in the previous chapter (pp.12,16) in connection with the convergent and discriminant validity of divergent thinking tests; therefore, only the findings relevant to the present discussion will be mentioned here.

Of the original group of 7th grade boys and girls tested in 1964, Cropley could only trace 111 for his followup study. This follow-up group was significantly higher in IQ and younger in age compared with the original group, but there was no significant difference in the divergent thinking scores of the two groups. Six Guilford type tests and IQ were used as predictors; the criteria of non-academic accomplishment were achievements in the field of art, drama, literature and music, derived from Wallach and Wing, with a slightly modified scoring scheme.

In view of the possible multi-dimensionality of the predictors and criteria canonical correlations were computed and it is reported that by adding IQ to the predictor battery the value of correlation was raised from .51 to .53 only. This Cropley takes as evidence of the long-range predictive validity of divergent thinking tests for non-academic accomplishments. But it was shown in the previous chapter that Cropley's divergent thinking tests and IQ have a higher average correlation than the divergent thinking tests themselves. If we add to this the further information that the average correlation of non-academic accomplishments with IQ and divergent thinking tests is .089 and .056 respectively, it seems that neither of these predictors contribute much to the variance of the non-academic accomplishments criteria.

It should also be pointed out that there were six different* divergent thinking predictors whereas the convergent abilities were represented by only one I.Q. index.

---

* The average correlation among these tests was .139 and .149 for boys and girls respectively as shown in Table II.2 in the previous chapter.

If other measures of convergent thinking had also been added to the predictor battery the contribution of this variable may have been greater. The use of canonical correlation procedure with such a predictor battery seems to have maximised the contribution of the divergent thinking measures. Even for the single IQ measure, Cropley does not report the canonical weight along with the weights given for other predictors and criteria. In view of these considerations we can conclude only that divergent thinking tests have reasonable validity for predicting non-academic accomplishment. The related conclusion regarding the lack of such validity in the IQ measure does not follow from the negligible increase in the value of canonical correlation when IQ is added to the predictor battery.

In Kogan and Pankove's (1972) study multiple regression analysis is used to indicate the relative contribution of productivity, uniqueness and IQ (at the fifth and tenth grade levels) to tenth grade non-academic accomplishments. Wallach and Wing's questionnaire was used in this study also. It will be recalled that in the follow-up study, Kogan and Pankove had administred divergent thinking tests under group and individual conditions in two sets of schools. The only significant contributions of any of the predictors to non-academic accomplishments was found to be in the schools where divergent thinking tests were administred individually. Even in these schools it was only fifth grade productivity which made the significant contribution. The significance levels for these three predictors are reported to be .10, .05 and .01 respectively. The first of these ($p < .10$) goes beyond the generally accepted level and so in fact only fifth grade IQ and tenth grade productivity are the significant contributors to non-academic accomplishments.

It should be noted that in terms of long-range prediction, it is IQ rather than either of the divergent thinking measures that makes the significant contribution.

Kogan and Pankove explain the generally negative findings in terms of the interaction of pupil anxiety with the impersonal atmosphere of the larger schools where divergent thinking tests were administred in group form. They suggest that in the larger school system "we should expect that extra-curricular activities pursued by test anxious subjects would have more to do with coping with fear of failure than with the cognitive dispositions tapped by creativity tasks" (p.439). To verify this they divided their subjects into high and low anxiety groups and correlated their ideational productivity scores with non-academic accomplishments. As predicted, the correlations for the low anxiety groups were higher than for the high anxiety group. From this analysis Kogan and Pankove conclude that "in the larger school system, validation of creative ability measures against out-of-classroom criteria of activity and accomplishment was obtained where anxiety (fear of failure) was not present as a disruptive element" (p.440). This sounds quite plausible, but, we have seen that even in the smaller school system no clear superiority of divergent measures over IQ as a predictor of non-academic accomplishment emerged.

That conventional aptitude/achievement tests and variables of a non-academic type are different in their predictive efficiency for forecasting non-academic accomplishments has also been shown in a study by Klein and Evans (1969).

Their sample consisted of over 2000 high school seniors in schools of different size and academic standing throughout the United States. These pupils had been given a number of standardised academic aptitude (SCAT) and achievement (STEP) tests when in the 7th grade. At the 9th grade level they also filled out a Background and Experience Questionnaire (BEQ) containing questions related to previous non-academic accomplishments. In the 12th grade the same pupils took the Independent Activities Questionnaire (IAQ). This questionnaire is very similar to the one used by Holland et al and covers accomplishments in such areas as science, art, design, music etc. It does however, differ from the earlier questionnaires in two ways. Firstly, it asks for details which could be used to verify a particular achievement claimed by a student. For example he has to mention where his creative writing was published or painting exhibited etc. Secondly, it is scored for quality or level of accomplishment rather than quantity.

In a correlational analysis of the above data Klein and Evans found that "7th grade academic achievement and ability measures predict 12th grade IAQ scores for "Science", "Writing", and "Arts and Science" at about .30. BEQ items and factor scores correlated with their corresponding IAQ scales even more highly than did the academic ability and achievement measures.... the combination of abilities, academic achievment and previous activities reliably predicted IAQ scores better than any of these measures alone" (p.153). In an earlier study with a college population, Skager, Shultz and Klein (1965) had also shown that if non-academic accomplishments are scored for quality rather than quantity, there is a significant correlation between IQ (SAT-V and M) and such accomplishments.

But in that study the only non-academic predictors used were socioeconomic background and an index of "Discussions", defined as "the number of hours spent in discussing topics such as scientific issues, world affairs, art, literature or drama with adults living in the house". Neither of these predictors are comparable to divergent thinking or past non-academic achievements type variables. So, although a positive relationship between conventional IQ and the quality of non-academic accomplishments has been shown in the Skager et al study, we do not know what the predictive validity of divergent thinking tests would have been in this case had they been used as predictors. What these two studies have shown is that conventional ability and achievement scores are not as irrelevant for predicting non-academic accomplishments as Holland et al and Wallach and Wing have argued.

Torrance, Tan and Allman (1970) did use the originality scores on divergent thinking tests as predictors of non-academic criteria in a longitudinal study with elementary school teachers. These criteria were based on a "Teachers Self-Inventory" covering such items of achievement as writing a story, organising a project, making-up an original dance, winning in a contest of creative work and so on. There were altogether 127 items in the Inventory to be checked on a five point scale. An empirical scoring key was developed after carrying out an item-analysis by comparing the responses of the top 27% on originality with those of the lower 27%. Of the 325 elementary education majors who had taken divergent thinking tests in 1958, 114 returned the Inventory in 1966. These inventories were scored using the empirically developed key.

Criterion scores derived in this way were correlated with originality and total creativity scores. Product-moment correlations of .62 and .57 were obtained with originality and creativity total respectively. From these findings Torrance et al conclude that "Teacher trainees identified as highly original in their thinking during their junior year appear to live more fully, be more fully involved in their teaching, and behave more creatively in the classroom than their less original counterparts...they appear to continue learning both in an independent fashion and in formal courses" (p.340 - 341). Thus, although the non-academic criteria used in this study are slightly different from the ones used by Holland et al the relationship of divergent thinking with these criteria is clear. To what extent there is an element of overlap with intelligence in this relationship cannot be said, as Torrance et al provide no information in this respect. If Wallach's (1970) analysis of most of Torrance's other studies is taken into account, it is quite possible that the divergent thinking measures used in this study also share a considerable proportion of their variance with IQ. Unfortunately, this missing information from the Torrance et al study makes their findings regarding the long term predictive validity of divergent thinking tests rather inconclusive.

In an earlier study, Hasan (1965) had also suggested that there may well be a positive relationship between divergent thinking ability and achievement in non-academic areas such as art, music, drama etc. This suggestion was based on the finding that teachers did rate pupils with high creativity scores as being more original and creative, although they gave the highest ratings for "desirability as a pupil"to those who were highest not only on VRQ but also on academic attainment in English and Arithmetic.

From this, Hasan reasoned that teachers' ratings for originality and creativity may be reflecting their awareness of the high creativity pupil's achievement in non-academic areas, whereas the ratings for "desirability as a pupil" were based on their recognition of the academic achievement of high VRQ pupils. However, this point was not pursued any further in the earlier study and was in fact a post hoc speculation regarding the obtained differences in teacher ratings for different qualities. In the present study the question of the validity of DT scores for predicting non-academic achievements will be considered in greater detail as the aims of the study stated in Chapter V require.

CHAPTER IV

THE ROLE OF MOTIVATION IN DIVERGENT THINKING AND NON-ACADEMIC ACHIEVEMENT

In Chapter I Guilford (1950) was quoted as saying that the empirical validity of divergent thinking tests cannot be considered in isolation from the question of personality and motivational factors which influence performance: "Creative productivity in everyday life is undoubtedly dependent upon primary traits other than abilities" (p.454). The same idea is implied in J.E. Anderson's (1960) discussion of the link between ability and achievement: "There are cut-off points or levels above which the demonstration of ability in relation to environmental demands is determined by the presence of other factors..." (p.24). Most studies of divergent thinking abilities have accepted the view that ability and personality factors interact to determine performance, and some have attempted to specify what these "primary traits" or "other factors" may be (Anderson 1966; Barron 1955, Di Scipio 1971a, 1971b; Garwood 1964; Getzels and Jackson 1962; Golann 1962; Hudson 1966, 1968; Leith 1972, Maddi 1965, Maw and Maw 1970; Shapiro 1966). Studies concerned with the personality characteristics of creative individuals (Mackinnon 1962, 1965; McClelland 1962, Roe 1952, 1953a; Taylor and Ellison 1964) are also relevant as they highlight some of the qualities these individuals have in common. Although these studies have not defined creativity in terms of divergent thinking scores, it can be seen that the personal qualities which characterise the creative scientists, artists, architects etc. have some similarity with the profiles of those who do well on divergent thinking tests.

Compare for example the following two statements - the first is from Mackinnon (1962), summarising his findings about outstanding architects; the second is from Getzels and Jackson (1963) describing their"high creatives" - i.e. pupils doing well on open-ended tests but not on conventional intelligence measures.

1. "But if I were to summarise what is most generally characteristic of the creative architect as we have seen him, it is his high level of effective intelligence, his openness to experience, his freedom from petty restraints and impoverishing inhibitions, his aesthetic sensitivity, his cognitive flexibility, his independence in thought and action...." (p.310).

2. "The high creatives tend to diverge from stereotyped meanings, to produce original fantasies, to perceive personal success by unconventional standards, to seek out careers that do not conform to what is expected of them" (p.202).

With some important reservations, Hudson (1966) also suggests that there may be a common personality dynamic underlying distinguished achievement and the convergent-divergent personality style: "Clearly it would be simple-minded error to confuse the answers to paper-and-pencil tests with writing a novel, say, or conducting a scientific experiment. Nevertheless, it may be that all elaborate and persistent thought has analogous origins... The habit of thinking, of pursuing ideas for their own sake, may be a by-product of the individual's need to keep the irrational elements of his personality under control" (pp.108-109).

Thus, although the studies in this area differ in the main focus of their inquiry and cover a vast ground from the divergent thinking ability of first grade children (Cartledge and Krauser 1963) to the psychological study of eminent scientists (Roe 1952), they all underline the crucial role of motivational factors in influencing the actual performance and achievement. Motivation in this context is used in a more general sense than the terms "achievement motivation" (derived usually from the TAT) or "academic motivation" (based on self-report questionnaires) imply. It is used to denote certain personality characteristics such as independence of judgement, preference for complexity, a disposition towards originality, driving absorption in work, risk-taking, curiosity, flexibility of attitude and an openness to experience, to mention but a few that have been studied. These qualities have often been found to be related to certain kinds of creative performance or achievement. For example, with reference to the threshold hypothesis discussed in Chapter II, it is argued that beyond the minimum IQ threshold, it is these motivational variables, rather than measured intelligence which determine performance.

Motivation and Divergent Thinking

One of the earliest studies to look at the correlates of a consistently original response pattern on open-ended tests is by Barron (1955). His subjects were a randomly selected group of 100 US Air Force captains. Originality was defined as the sum of statistically rare or unique responses (within the limits of appropriateness and "adaptive to reality") given to a number of tests. Some of these tests (Unusual Uses, Consequences, Plot Titles) are now regularly used as part of a divergent thinking test battery.

On the basis of these test scores Barron identified two groups of subjects, referred to as "regularly original" and "regularly unoriginal" and hypothesised certain characteristics which would be expected in the former but not in the latter group. In the light of previous research regarding conformity, yielding etc. Barron reasoned that the original persons:

1. Would prefer complexity and apparent imbalance in phenomena.
2. Are more complex psychodynamically and have greater personal scope.
3. Are more independent in their judgements.
4. Are more self-assertive and dominant.
5. Reject suppression as a mechanism for the control of impulse.

Using various tests, inventories and ratings (Barron-Welsh Figure Preference Test, California Personality Inventory etc.) as measures of the hypothesized dependent variables, Barron found that there was a significant difference ($p < .05$ or better) in the expected direction on twelve out of the fifteen measures he had used. That is, the regularly original group were found to have a preference for complexity, to be more complex persons, to have independence of judgement, to be assertive and to reject suppression of impulse. Using essentially the same research design as Barron and very similar measures of the independent and dependent variables, Garwood (1964) also found that the high creatives were significantly higher on cognitive flexibility, dominance sociability, social presence and self-acceptance. They also showed greater integration of nonconscious material. They were less conforming and lower on self-control and the desire to make a good impression.

So here we have some indication of the personality characteristics of people who respond with original ideas when given open-ended tests. Given the above characteristics, it is possible that they will be better at tasks which call for these qualities rather than those which require an opposite personal style. In regard to the suggested link between divergent thinking and non-academic accomplishments (Wallach and Wing 1969) it is these motivational variables which may be mediating between ability and performance.

A consideration of some other relevant studies in this field seems to support the above view. For example, inspite of the fact that Getzels and Jackson (1962) did not find a significant difference in the achievement motivation scores of their highly intelligent and highly creative groups, they do report significant differences in their values, aspirations, fantasies and family background. It has been mentioned before that Getzels and Jackson ascribe the superior academic achievement of the high creatives to their cognitive giftedness. But it may also be that although the values and aspirations of the high creatives are different from those of the highly intelligent group, these values and aspirations are more closely related to academic achievement than are creativity scores. In view of Wallach and Kogan's (1965) arguement that Getzels and Jackson's creativity battery is not clearly distinguishable from conventional intelligence measures, it is difficult to see how scores derived from such a battery can be making any independent contribution to achievement. In these circumstances a motivational explanation (in the wider sense of the word, not just in terms of TAT scores) of the equally superior achievement of the high creatives seems more convincing.

In view of the consistently different personality descriptions of subjects scoring at the extremes of the Barron-Welsh Art Scale (Barron 1952, 1958; Barron and Welsh 1952, Rosen 1955), and the Revised Art Scale (Welsh 1959b) of the Welsh Figure Preference Test, Golann (1962) has proposed a "creativity motive" behind these differences in personal style. He defines this motive as referring to "the tendency for individuals to differ in the degree to which they attempt to experience their fullest perceptual, cognitive, and expressive potentials in their interaction with their environment. This will often lead to behaviour that is creative in terms of the individual's previous repertory, and occasionally to behaviour which is judged by others to be creative in a larger sense" (p.590).

Golann linked this creativity motive construct to preferences on the Art Scales (Barron, Welsh - BW and Revised Art Scale - RA) by reasoning that "individuals strongly motivated to experience their perceptual cognitive and expressive potentials would prefer objects and situations which permitted more idiosyncratic ways of dealing with them" (p.591). From this he further hypothesized that the figures in the RA scale which were preferred by artists should be more "ambiguous or evocative" than the figures disliked by them. Ambiguity was defined as the quality of a figure to call forth "a multiplicity of different associations from a group of Ss, it does not refer to qualities of haziness, lack of structure, or any other apriori stimulus considerations." Having operationalised the concept of ambiguity in this way, Golann did find that figures on the RA Scale which had been preferred by artists, who presumably were high creativity motive subjects, were indeed more ambiguous than the figures not liked by them.

In a second experiment, Golann extended his hypothesis to state that "individuals who prefer the ambiguous, evocative, dynamic WFPT* stimuli, prefer activities which are more allowing of self-expression while Ss who prefer the simple, static figures prefer more routine structured activities" (p.593). This study was carried out with sixth and eighth grade pupils who completed the WFPT and a 36 item questionnaire regarding the sort of activities they preferred to pursue. The questionnaire was forced-choice in format, consisting of two alternatives for each item. On rational ground the alternatives were chosen in such a way that one referred to an activity of a creative type allowing for self-expression (i.e. drawing a picture) and the other to more convergent type of activities (i.e. working on a jigsaw puzzle, colouring in a colouring book etc.). A score of one point was earned when the "creative" alternative was chosen, no credit was given for the other alternative. Test-retest reliability over a three week interval is reported to be .66 and .86 for samples of sixth and eighth grade pupils, respectively.

The WFPT and questionnaire scores were analysed in two ways. Firstly, correlations were obtained between preference for an ambiguous figure and creative activity, for all eight classes separately. With one exception (- .14, $p > .05$) all correlations were positive ranging from .16 to .41. Four of these seven correlations were significant at the 5% level, the remaining three were non-significant. Secondly, groups of subjects falling above and below one standard deviation on the WFPT were compared for their choice of alternative activities on the questionnaire items.

---

*WFPT refers to the Welsh Figure Preference Test from which the Barron-Welsh Art Scale (BW) and the Revised Art Scale (RA) are derived.

On 27 of the 36 items, those preferring ambiguous WFPT figures also had a higher percentage of creative activities on the questionnaire. However only eight of these percentage differences were significant ($p < .05$). It should be noted that this study was carried out only with boys and Golann reports that in another study (Golann 1961) results for girls were even more mixed. It is not clear whether the same questionnaire items were used with girls also or different ones. If they were the same, then it is possible that in view of the differential socialisation of girls and boys in terms of which activities are appropriate for them, the questionnaire may not have been suitable for girls.

Although Golann's findings are not conclusive in the sense that only a small number of significant correlations were obtained and the questionnaire items which discriminated beyond chance were also small in number (8 out of 36), it does suggest that activities which allow for self-expression and initiative appeal to people who are "tolerant of ambiguity", or indeed prefer it to more structured stimuli/situations. The relationship suggested in Golann's study provides some empirical support to the argument presented in Chapter III regarding the differential validity of various predictors only for criteria which are conceptually relevant. With reference to the specific question of the prediction of non-academic accomplishments, it seems that preferences of the type revealed by the WFPT may be as important as scores on divergent thinking tests.

Maddi (1965) has shown that scores for originality on the Uses Test have a correlation of .57 with a Novelty Score derived from TAT/

TAT stories*. The ability to give unusual endings to a Similies Preference test has also a correlation of .45 with the Novelty Score. Maddi explains these relationships in terms of a "need for novelty as part of [the] motivational complex". He found that the Novelty Score also had positive correlations with otherindices of preference for novelty as a personality characteristic. For example, on Cattell's 16PF the correlation was .28 with "experimenting tendency" and on the Stern Activities Index it was .25 with "need for change". So it seems as if the cognitive ability to think of original uses for objects, or unusual endings for incomplete similies, is itself related to a personal style which consistently interacts with its environment in a novel and unusual way. Therefore, any consideration of divergent thinking tests as predictors of non-academic accomplishments implies that motivational variables may also be affecting the obtained correlations.

In view of the generally positive correlations found between convergent and divergent thinking tests, Anderson (1966) has suggested that a study of the non-intellective correlates of originality may provide better evidence for the construct validity of divergent thinking tests. The theory on which he based his empirical work is derived from Luria's ideas regarding the mechanism of self-regulation and self-reinforcement by internalising adult instructions. Anderson reasoned that "the original individual has imperfectly internalised parental rules" and therefore is more likely to be impulsive and non-conforming.

---

*The scoring scheme used by Maddi was entirely different from the one devised by McClelland (1953) and generally used in analysing TAT protocols for achievement motivation. Actually, Maddi reports that the Novelty Score <u>does not</u> correlate with the nAch score also obtained in the same study.

Talkativeness may also be a correlate of originality, as "the generation of large quantities of words may allow the individual to capture more easily unconscious material" (p.287). To test these ideas, Anderson used the scores on Consequences and Tin Can Uses as a criterion of originality. The predictors were Vocabulary, Verbal and Non-verbal IQ, as well as measures of non-intellective variables such as socio-economic status (SES), impulse expression, risk-taking etc. From a multiple regression analysis Anderson concluded that "the correlates [of originality] most powerful in terms of the extent of their contribution to the variance were the convergent tests. Also making a significant, though much less powerful, contribution was impulse expression, while the non-intellective tests made negligible contribution" (p.289). When convergent measures were removed statistically from the analysis, variables such as SES and risk-taking did begin to make a significant contribution to originality, but it is interesting to note that impulse expression did not contribute significantly in this analysis. Thus, Anderson's idea that impulsivity may be one of the non-intellective correlates of originality finds some support from the first mode of analysis. At the same time this study further confirms what others have reported before - that there is a considerable overlap between measured intelligence and scores on divergent thinking tests.

Another study which has looked at the possible relationship between risk-taking and divergent thinking ability is that of Pankove and Kogan (1968). Out of three risk-taking measures only one was found to be related to scores on two of Wallach and Kogan tasks, and that also only for boys.

IQ was also found to be related to the same risk-taking measure and so no clear association between divergent thinking and risk-taking emerged as had been hypothesized. Pankove and Kogan consider the possibility of the moderating effects of anxiety as an explanation for these mixed findings. But when they carry out an analysis with different aspects of this variable (General Anxiety, Test Anxiety and Defensiveness), again the only significant effects were for boys on the shuffleboard task. That is, high Test Anxious boys were low on creativity, and low Defensive boys were high on creativity. There were no significant effects for girls on any of the three tasks and also none for the boys on the remaining two risk-taking tasks.

Shaefer and Anastasi (1968) noted that most research in the field of creativity had been based on tests of one kind or another and this approach has usually not taken into account motivational factors which they consider are of equal, if not greater importance. These authors cross-validated a biographical inventory for identifying creativity in adolescent boys and found various correlates of artistic and scientific creativity. The main criterion of creativity was teachers' nomination, "supported by specific creative products". A s a check Guilford's Alternate Uses test was also given. Validity coefficients in the cross-validation sample were .64 and .35 for the artistic and scientific keys respectively. Final keys containing items differentiating in the initial and cross-validation samples at $p<.05$ were used to describe the similarities and differences in the motivational characteristics of creative students.

Both, in the field of arts and science creative students

1. Came from families which not only gave "academic support" but also provided role models of interest and creative expression in the student's own field.

2. Aspired for Ivy League Colleges and were less interested in sports. Shaefer and Anastasi describe them as having "a strong intellectual and cultural orientation".

3. Had a "pervasive and continuing enthusiasm" for their chosen field and pursued it with a single-mindedness of purpose.

4. Had a breadth of interest, i.e. they checked more subjects as favourite in HS* than their controls.

5. Showed a stronger drive towards novelty and diversity.

Specific differences between the background and experiences of creatives in the arts and science are also reported. For example, those showing creativity in the arts had "greater exposure to environmental diversity" - their parents had travelled and moved more, creative scientists had more conventional parents and they themselves had a stronger sex role identification. The artistic creativity group also reported more day-dreaming and less social participation.

These descriptions of the creative group are generally in keeping with the findings of earlier studies (Barron 1963; Getzels and Jackson 1962; Hudson 1966, 1968, Wallach and Kogan 1965).

* High School.

But it is not clear from Shaefer and Anastasi's study, to what extent these differences could also be explained in terms of IQ. For example, such characteristics as aspiration for Ivy League Colleges, or a liking for a number of school subjects, or a "strong intellectual and cultural orientation" may very well characterize high IQ pupils too. The authors did not consider IQ at all in this study, although they do report matching the creatives and controls for GPA.* It is true that GPA* and IQ have generally been found to be positively correlated, but the correlations at best are around .7. This accounts for only half of the identifiable variance and is a very indirect way of controlling for IQ. The study would have been more definitive had the role of IQ been taken into account explicity.

Another study which concluded that high curiosity boys have higher creativity scores is that by Maw and Maw (1970), but these authors also do not consider the concomitant contribution of IQ to the obtained differences. Creativity was measured by the Word Association Test and boys were divided into high and low curiosity groups on the basis of teacher and peer ratings. A number of personality tests and questionnaires were also given (California Test of Personality, Children's Personality Questionnaire etc.), and several creativity factors were identified for the high and low curiosity groups. High curiosity boys were found to be at the positive end of Restrained Creativity factor and Impulsive Creativity factor. According to the authors, "Restrained Creativity is a factor that is loaded positively for dependability, efficiency, promptness, self-control" (p.328).

*Grade Point Average

For Impulsive Creativity the authors state that "One high on this factor is intelligent and consistent in his thinking, but unrealistic about his own abilities", and they conclude:"This study has indicated that boys who differ in curiosity also differ in creativity" (p.329).

As mentioned before, differences in the IQ of high and low curiosity boys have not been taken into account. However, from the factor loadings reported it can be seen that the high curiosity boys showed a significant ($p < .01$) positive loading (.425) for creativity on the first factor, identified as the General Factor. It should be pointed out that the verbal battery of the Lorge-Thorndike Intelligence Tests had also been administered to all the boys and it would presumably have a positive loading on the General Factor too, although this is not reported (only the creativity loadings on the various factors are reported). Another point to be noted is that the measure of creativity used was the Word Associations test, which, has generally been found to have higher correlations with conventional intelligence, than some of the other divergent thinking tests (Getzels and Jackson 1962, Hasan and Butcher 1966). Indeed, Hudson (1966) has remarked that it is not a truly open-ended test. In view of these considerations, the relationship between curiosity and creativity postulated by Maw and Maw is not at all clear from this study.

So far studies of the relationship between motivational factors and divergent thinking ability have been inconclusive probably because they have attempted to study such complex variables as anxiety, risk-taking, curiosity etc.

Ward and Kogan and Pankove (1972) consider that under these circumstances it may be better to study "somewhat simpler motivational variations". For this reason they undertook a study "to test the effect of a concrete incentive on children's ideational production". In this study their main concern was with two issues: Firstly, they wanted to see if the relative individual differences in divergent thinking scores were stable over different testing conditions, some of them offering a monetary reward for producing ideas. Secondly, they also looked at the related question of whether individual differences in DT scores should be "interpreted primarily as differences in the capacity for such productions or as differences in motivation". Obviously, the two questions are linked in the sense that if performance is stable under different motivational contexts the role of ability rather than motivation will be considered to be more crucial and vice-versa.

The details of Ward et al's study have been mentioned earlier in Chapter II (p. 23) while discussing the convergent and discriminant validity of divergent thinking tests, so only those findings relevant to the motivational issue will be considered here. One verbal and one non-verbal Wallach and Kogan (1965) test was given to all subjects under the "baseline" condition of individual, untimed, non-evaluative administration. Then one third of the subjects of each sex were assigned randomly to three experimental groups - No Reward, Delayed Reward (one US penny for each answer but given at the end of the testing session) and Immediate Reward (penny given as each idea produced).

Again, another set of verbal and non-verbal tests was given to all the subjects under these different reward conditions. Data obtained in this way were subjected to an analysis of covariance, with scores under the incentive conditions as the dependent variable and the baseline score as the covariate. Ward et al report that the different reward conditions "left unaffected the magnitude of the performance difference between subjects who gave many ideas under baseline conditions and those who gave few" (p.673). Thus, in answer to the first question posed at the beginning of their study, evidence for stability of the initial individual differences under different motivational contexts was obtained. Indirectly, this also provides an answer to the second question regarding the greater importance of ability or motivation in performance on divergent thinking tests. Ward et al conclude that although the Immediate and Delayed Reward treatment did succeed in significantly increasing the number of ideas produced, "these treatments failed to diminish the difference in ideational production between those who produced few ideas under baseline conditions and those who produced many, indicating that variation in motivational level is not sufficient to account for this difference. The alternative explanation - that the observed individual differences reflect variation in the capacity for divergent ideational production - thus receives support" (p.675).

Motivation and Non-academic Accomplishment

Although the studies considered above are far from conclusive they do suggest that to some extent people are consistent in their likes and dislikes of certain kinds of achievement (Getzels and Jackson 1962, Golann 1962, Shaefer and Anastasi 1968) and in their ability to give original responses to certain cognitive tasks (Barron/

(Barron 1955, Maddi 1965, Anderson 1966, Maw and Maw 1971). The study by Ward et al (1972) has further shown that a higher level of performance under conditions of increased motivation is not in itself sufficient to show that motivation rather than ability is the more crucial factor in performance, unless it can be shown at the same time that incentive conditions also altered the relative rank order of individuals obtained under "baseline" conditions. From this it is possible to argue that the desire to do well in a preferred sphere of achievement (this is what motivation means) may itself be linked with the perception of how good one is at it (and this is what ability means). Hudson (1968) has strongly argued for such a moderating effect of self-perception between ability and achievement: "It seems that we are bound to envisage the intellectual growth of the individual, the evolution of his characteristic frame of mind, as the product not only of his genetic endowment and hormonal secretions, but of a continual traffic with his context - with parents and teachers, examinations and curricula, prejudices and myths. And even when his frame of mind is firmly established, it seems that an individual's intellectual performance is partially conditioned by the audience for which, and the setting in which, it is produced" (p.103).

On this view the one general weakness of the studies regarding the role of motivational variables in divergent thinking is their failure to take into account the situational and social determinants of motivation. Katz (1967) has made this criticism with reference to the concept of achievement motivation in particular as proposed by McClelland et al (1953).

But it has relevance for the studies under consideration in this chapter also. The main point raised by Katz is that achievement motivation is too general and ambiguous a concept. It fails to take into account the nature of the task, and its appeal to the individual, for which motivation may or may not exist. Katz illustrates this point by quoting the typical finding of low achievement motivation in working class pupils and argues that it does not necessarily denote a lack of the desire to achieve. Rather, it may be that the sort of achievement valued by this group is not congruent with the generally accepted definition of achievement in the school setting. He suggests that it would be better to think in terms of specific motives for specific kinds of achievement rather than a general concept of achievement motivation which has now become almost synonymous with academic motivation.

In the alternative approach suggested by Katz we see a link with the arguments for broadening the predictors and criteria of achievement put forward by Getzels and Jackson (1963), Holland et al (1962, 1964, 1965) and Wallach and Wing (1969). For example, if different tasks or achievements have a different appeal for individuals, as Katz has suggested, then recognising non-academic criteria of achievement as legitimate and valuable is one of the obvious ways of providing an opportunity to the non-academic student to show where his interests and strength lie. In fact, throughout their book, Wallach and Wing imply a motivational explanation for the accomplishments of the non-academic achiever: "Extracurricular attainments... tell us something about what a student undertakes in that sector of his life where - unlike the case with academic achievement - he is, by and large, doing what he wants to do rather than what society requires of him.

We suspected that self-initiated accomplishments would be more representative of a persons characteristic functioning under the real life conditions he encounters when formal schooling, is over than would academic grades" (Preface VI). And again, writing about the implifications of their findings for education they say "our orientation in the present work has been toward delineating those students who most clearly are running on their own power in the world - the students who are most strongly committed to particular lines of endeavour which are carried on for reasons intrinsic to the tasks themselves" (p.129).

Inspite of this emphasis on the motivational determinants of non-academic accomplishments, whatever evidence Wallach and Wing provide in the book is in terms of "high ideational productivity", "ideational resourcefulness" "cognitive vitality or energy" and "thinking capacity" etc. The discussion of their findings (pp.72-80) has shown that they do not systematically take into account the motivational factors which they consider so crucial for non-academic accomplishments. In his review of Wallach and Wing's book, Nicholls (1970) has pointed out that "ability and motivation are almost certainly interdependently involved. However, a comparative test of motivational and ability interpretations could be made and would help clarify the nature of the phenomena underlying these interesting results" (p.278). Nicholls goes on to offer some motivational interpretations on theoretical grounds. He argues that the validity of divergent thinking tests depends to a large extent on the "similarity of testing and criterion conditions". This is supported by findings quoted from Datta's (1963) study in which "scientific creativity in industry was predicted when instructions emphasized the need for original and worthwhile responses, but not when instructions were neutral". Nicholls concludes that Wallach and Wing's findings can be explained in a similar manner: "A motivational explanation of Wallach and Wing's results is that the test, as they administred them, measure a tendency to develop task-relevant motivation where extrinsic supports or pressures and constraints on mode of task performance are minimal"* (p.278). It is surprising that the two partial replications of Wallach and Wing's study (Cropley 1972, Kogan and Pankove 1972) have also not looked for any direct empirical evidence for the role of motivational variables in non-academic accomplishments. As the next chapter shows, one of the aims of the present study was to see what evidence could be found for a motivational explanation.

---

*Wallach and Wing had obtained their data on divergent thinking tests and the non-academic accomplishments questionnaire by mailing these instruments to prospective Duke University freshmen during the summer holidays. The students were requested to complete the tests and questionnaires at home if they were willing to volunteer for the study.

SECTION TWO

DESCRIPTION OF THE PRESENT STUDY

CHAPTER V

AIMS AND PLAN OF RESEARCH

Aims

This study was undertaken to look into the following three questions related to the "threshold" or "ability gradient" theory of the relationship between IQ and divergent thinking, discussed in Chapter II:

(a) Does the relationship between measured intelligence and divergent thinking become weaker as we move up the intelligence scale, until the two variables become independent of each other around IQ 120, as supporters of the threshold theory believe (Barron 1963, Mackinnon 1962, McNemar 1964, Torrance 1960, 1962)? The review of literature in the previous chapters indicates that so far the evidence in favour of this hypothesis is equivocal (Ginsburg and Whittemore 1968, Haddon and Lytton 1968, 1971, Lytton and Cotton 1969, Yamamoto 1965). Also the very basis of a long standing debate about the existence of two different cognitive styles (Burt 1962, Cronbach 1968, Getzels and Jackson 1962, Guilford 1959a, Heim 1970, Hudson 1966, 1968, Wallach and Kogan 1965) may become clearer if the threshold theory can be supported by empirical evidence. Then it may be said that over a whole range of IQ there is a significant relationship between intelligence and divergent thinking, but at higher levels of IQ this pattern breaks down.

(b) The next question, following from the one above is about the relative contribution of IQ and divergent thinking to actual achievement in the academic and non-academic fields of endeavour.

An extension of the threshold hypothesis is that beyond a certain level of IQ (again set around 120) divergent thinking begins to have a significant effect on performance and the effect of IQ declines (Anderson 1960, Holland 1961, Hoyt 1965, Hudson 1964b, Richards et al 1967, Wallach and Wing 1969). Again the evidence for such an interaction is either equivocal (Clarke, Veldman and Thorpe 1965; Cline, Richards, Abe 1962; Yamamoto 1961, 1964a,b,c) or it fails to support the hypothesis (Cicirelli 1965; Edwards and Tyler 1965; Flescher 1963; Haddon and Lytton 1971).

(c) The third question, related to the above two is about the effect of motivational variables on performance. It is argued that beyond the IQ threshold of 120, motivational factors also begin to have a significant effect on performance and differences in measured intelligence become less crucial (Barron 1963, Butcher 1968, Hudson 1964 b , Maddi 1965, Torrance 1963, Wallach and Wing 1969). In view of the work of Holland et al (1964, 1965, 1967) and of Wallach and Wing (1969) in this field, the criterion of achievement in the present study has been widened to include academic as well as non-academic achievements. Specifically the question may be put as follows: is there a significant relationship between achievement motivation (assessed by the method of content analysis of imaginative stories) and performance in the academic and non-academic spheres and especially beyond the IQ level at which performance begins to vary independently of IQ? However, evidence based on the TAT is indirect in the sense that it is derived from the <u>interpretation</u> of a story.

It was therefore decided to look at more direct evidence of motivation in the form of career aspirations given by the subjects to the Careers Officer, their choice of subjects* and their leisure time activities.

Plan of Research

The School:

The study reported in this thesis was carried out at a large (enrolment 1800+), new, local authority comprehensive school in Scotland. The catchment area for the school consists mainly of council housing estates and the socio-economic composition of the pupil population was described by the Deputy Headmaster of the school as "weighted towards the lower end of the scale". In terms of ability, the pupils at the school may be considered representative of the general population of local authority comprehensive schools, except for a very slight "creaming off" of pupils from the top end of the ability continuum to attend a high prestige local authority school**.

---

*For a discussion of how subject choice may be an index of the level of aspiration, see pp.111-116 in this Chapter.

**This very brief description of the school seems necessary here in view of the many fee paying schools in Edinburgh which "cream off" pupils from local authority schools, thus creating considerable differences not only in the socio-economic composition of the school population but in the ability range also.

The Pupils:

The aims of the research required that pupils from the entire range of ability be included in the study. At the same time, in order to test the "threshold" hypothesis it was necessary to have more pupils at the higher levels of ability so that groups of reasonable size might be formed between the different cut-off points on the intelligence scale.

The pupils tested were in their third and fourth year of secondary education and were grouped into Certificate classes and Leavers; the former being those who had chosen to stay on beyond the statutory leaving age to take "O" grade or Highers, the latter were going to leave school at fifteen. For the present study, four Certificate classes and one Leavers' class were chosen. In the Certificate group an attempt was made to include classes at the highest and middle levels of ability. The Certificate classes in this school are grouped into "sets" according to their performance in English in the school examinations. Two classes from the top set and two from the middle set were selected for testing.

Girls and boys were in roughly equal numbers (girls 53, boys 55). The mean IQ (on Moray House Intelligence Test, (Adult)I) of the whole group was 110.296 with a standard deviation of 11.264. Means and standard deviations of the different classes as well as of boys/girls, third year/fourth year, Certificate classes/Leavers are given in Appendix A, Tables 1.1 to 1.6

Groups Studied:

<u>School Classes</u>: It can be seen from Table 1.1 that in the Certificate group the significant difference in IQ is between school sets (1, 4) rather than between third and fourth year classes.

Therefore for the purpose of this study the two top sets from the third and fourth year (4C1 and 3C1) have been combined. Similarly, the two middle sets (4C4 and 3C4) are treated as one group and the Leavers' class which has a significantly lower IQ than the four Certificate classes forms a group of its own. There are two main reasons for including the Leavers'class:

(a) As mentioned before, one of the aims of the present study was to find out what contribution, if any, achievement motivation makes to an individual's performance in the academic and non-academic fields. The Leavers' class was a group which had chosen to leave school at the earliest opportunity and was therefore not going to attempt to get any formal academic qualifications such as "O" grades or Highers by staying on at school beyond the minimum leaving age*. If the variable of achievement motivation has any predictive validity, the Leavers' class would be expected to form a low achievement group in this context and would have a low motivation score in comparison with the Certificate classes.

(b) Since an important aspect of this study was to look for empirical evidence for the threshold hypothesis of a decreasing relationship between intelligence and divergent thinking ability, as the level of intelligence rises, the Leavers' class also provided a group of pupils in the lower ranges of intelligence for comparison with the higher intelligence groups.

---

*Most of them do however get a local school certificate based on examinations at the end of the third year. It is generally known by the pupils that this Certificate is much lower in its "value" to "O" grades and Highers. Of course, some of the leavers may go on to get academic qualifications through evening classes or day release.

Thus we have three groups of pupils differentiated in terms of ability, as well as on the basis of subject choice and over all status within the pupil heirarchy, as the discussion below will show.

Strictly speaking the school does not stream pupils. Since English and Arithmetic are compulsory subjects for all Certificate and non-Certificate classes, the same sets which are formed on the basis of English attainment also go to Arithmetic lessons as a group. Apart from this, pupils from different English sets can theoretically, mix for other lessons, depending on their choice of optionals, at least within the Certificate group. In practice, however, there seems to be a "tendency" for the top English sets to opt for "difficult" subjects like Maths, Physics, Chemistry, Latin and German. This trend is equally present for boys and girls but in the middle sets girls go in more for applied subjects like Secretarial Studies, Commercial Studies, Dress and Design and Home-Management, while boys opt for Technical Drawing, and Mechanics as the following tables show.*

---

*This information about subject choice was obtained from school records.

Table V.1 : Number and Percentage of Pupils in the Top and Middle Sets, taking "Difficult" Subjects. Boys and Girls, n = 82.

Subject	Top Sets n=43		Middle Sets n=39	
	no.	% age	no.	% age
Art	7	16	1	3
Chemistry	24	56	17	44
German	11	26	1	3
History	13	30	9	23
Latin	3	7	0	-
Maths	39	88	25	64
Physics	18	37	11	28

Table V .2 : Number and Percentage of Boys in the Top and Middle Sets taking "Easy" subjects. n = 38

Subject	Top Sets n=21		Middle Sets n=17	
	no.	% age	no.	% age
Mechanics	8	38	13	76
Tech. Drawing	10	48	14	82

Table V .3 : Number and Percentage of Girls in the Top and Middle Sets Taking "Easy" Subjects. n = 44.

Subject	Top Sets n=22		Middle Sets n=22	
	no.	% age	no.	% age
Commercial Studies	0	-	12	55
Dress and Design/Home-Management	0	-	5	23
Secretarial Studies	2	9	14	64

Looking at these tables* it is clear that the IQ difference between the two** groups is associated with systematic differences in subject choice as well. In a study of schoolchildren's perception of subjects Cowan (1971) has shown that children perceive differences in subjects in terms of the (1) theoretical-practical, (2) masculine-feminine, (3) science-arts, and (4) relevance to vocational goals - nonrelevance to vocational goals, distinctions. In the present study Table V.1 may be differentiated from Tables V .2 and V.3 in terms of (1) and (4) above, The theoretical-practical distinction is obvious, the relevance to occupational goals can be seen if the choice in Table V.1 is considered as one leading to higher education and that in Tables V.2 and V.3 as that leading to relatively more immediate occupational goals.

It is possible that the above differences in subject choice contribute to the building up of a group identity amongst pupils within the context of the school where subject hierarchies are at least implicitly known and choosing a practical (i.e. easy) rather than a theoretical (i.e. difficult) subject places the group making such a choice somewhere below the highest level.

---

*The difference in percentages is so much that it seems unnecessary to apply a statistical test to establish that the differences could not have arisen by sheer chance. Indeed, it would be difficult to find an appropriate test in this case as the categories in the tables are over-lapping ones and the data are such that neither the column nor the row totals add up to 100.

**The Leavers class was excluded from this analysis as the options open to them are different from those of Certificate classes and therefore the comparison would not have been legitimate.

In other words, although there is no streaming in the school, the top English sets seem to have an "A stream" and the middle sets a "B stream" identity.*

The purpose of dividing pupils into Top, Middle and Leavers' sets was to examine the threshold hypothesis in a situation where differences in IQ are accompanied by evidence of differential over all standing of the groups within the pupil heirarchy, which is likely to influence their perception and performance. That is, the pupils' definition of the situation in terms of what is expected of them as a member of the Top, Middle or Leavers' set, may affect their divergent thinking score and its relationship with IQ as much as the level of IQ itself. So far, most studies of the threshold hypothesis (Cicirelli 1965, Ginsburg and Whittemore 1968, Haddon and Lytton 1968, Lytton and Cotton 1969 and Yamamoto 1965) have taken a wide ability range from <u>different</u> classes, put them together and then studied the IQ-DT correlations at different cut-off points on the IQ scale. Such an experimental design assumes complete independence of IQ and divergent thinking from such variables as classroom climate and pupils' perception and interpretation of their own position in the complex organisation of the school. Studies by Hargreaves (1967), Keddie (1971) and Lacey (1970) suggest that these contextual variables may have a subtle but significant influence on performance.

---

*Hargreaves (1967), Keddie (1971) Lacey (1970) point to the existence of an image of the typical A, B or C stream pupil among teachers and children alike, based on assumptions about ability, behaviour, aspirations and home background etc.

More direct evidence on this question is also available in research undertaken specifically to study the effect of school atmosphere or testing conditions on performance on open-ended tests. Boersma and O'Bryan (1968), Haddon and Lytton (1968), Madaus (1967), Ogletree (1971), Torrance (1964) and Vernon (1971) all found that informal and relaxed conditions yield higher DT scores, but in a cross-cultural study Marino (1971) reports more equivocal findings. Although these studies have shown the effect of motivational and environmental factors in divergent thinking, they do not provide any information on the interaction of IQ level with these variables.* On the other hand, it may be that studies of the threshold theory that have combined several classes before obtaining their sub-groups at different IQ levels have obtained equivocal results both in terms of IQ-DT relationship and in terms of the additional contribution of DT to school attainment beyond IQ115-120, precisely because they have invariably considered level of IQ alone as the crucial variable, regardless of the contextual factors mentioned earlier.

It may of course be objected that by studying ability groups as proposed here one cannot really find answers to questions about the threshold hypothesis since it is quite possible, and often likely, that although the mean IQ of the groups may differ significantly, there may be individuals in both groups who have the same IQ.

---

*Haddon and Lytton's (1968) study discussed in the previous chapter is an exception in this respect.

If this is the case then clearly the overlap in IQ would seem to render such a research design meaningless for the threshold hypothesis. However, the overlap situation itself provides a way out of this difficulty. For example, it is possible to compare the IQ/DT relationship for the initially different ability groups by matching them on IQ so that the grouping in terms of school classes or whatever other differentiating factors are operative, become crucial for explaining the difference in IQ/DT correlations.

As mentioned before in English sets in this school are not formed on the basis of IQ as such and the concomitant sorting out of the Top and Middle sets in terms of significant mean difference in IQ as well (Table 1.4, Appendix A) could have been a consequence of the high correlation between IQ and English attainment ($r = .679$, Table 2.1, Appendix A). Since the setting is not according to IQ it was anticipated that there would be considerable overlap in individual IQs between the two groups. According to the procedure suggested above, it was decided to match individual pupils from the Top and Middle sets on IQ and then compute IQ/DT correlations for them to find out if, despite the similarity in IQ, the correlation between IQ and DT remains different in the same direction as when all the pupils in the Top and Middle sets are included in the correlational analysis.

<u>IQ Groups</u>:

Another way of dealing with the restriction of range problem is to divide up the group in such a way that the subgroups have equal standard deviations.

T.L. Kelley (1947) has provided a method whereby making use of the known mathematical properties of the unit normal distribution, cut-off points on a normally distributed variable may be worked out in such a way that the dispersion of scores at the different levels is very similar. One main disadvantage of this procedure is that given the shape of a normal distribution, we end up with the majority of cases at the intermediate level/levels and very few at the extremes. This however is only an apparent shortcoming as the homogeniety of variance at the different levels makes it possible for us to establish whether or not, inspite of the smallness of the extreme groups, a particular difference is significant beyond chance. Using Kelley's method* the entire sample in the present study was divided up into three equal variance groups referred to as the High, Middle and Low IQ levels (Table 1.5, Appendix A).

---

*Kelley's (1947) formulae for working out the mean and standard deviation of a portion or a "slice" of a normal distribution are given in Appendix C. A computer programme in IMP language was written to try out various possible cut-off points that were likely to yield very similar standard deviations in a 3-level split on IQ. Although the mean IQ was 110.296 which is above the norm mean of 100, it was normally distributed and therefore Kelley's method could be applied to the data.

Experimental Groups:

So far two ways of grouping pupils for the present study have been described - grouping by school classes and grouping by a statistical procedure so that the scores have similar spread at each level of IQ. It was thought that the relationship between the independent variables (IQ, Divergent Thinking, Achievement, Motivation and Social Class) and dependent variables (School attainment, Non-academic Accomplishments and Career Choice) could be studied in greater depth if experimental groups are isolated on the basis of IQ and DT and then their standing on the dependent variables is compared using an analysis of variance technique. The terms "high" and "low" were defined in the following way.

Pupils from the top third of the IQ and DT distribution were identified separately. Thus there were 34 pupils in the IQ group and 34 in the DT group. In view of a significant positive correlation between these two variables ($r = .467$, Table 2.1, Appendix A) it was inevitable that some individuals would be present in both these lists. This is the group that will be referred to as High-High in the rest of the present study. As it happened there were exactly half the pupils (i.e. 17) in each of the IQ and DT groups who fell into this category. This left 17 in the group identified as High-IQ (those in the top third of the IQ distribution but <u>not</u> in the top third of DT) and also 17 in the High-DT group (i.e. those in the top third of DT, but <u>not</u> in the top third of IQ). A fourth

Low-Low group was formed by identifying 17 pupils falling in the lowest third of the IQ and DT distributions. Table 1.6 in Appendix A gives the mean scores and standard deviations of these groups along with that of the remaining group for purposes of comparison.

Administration of Tests and Questionnaires

The following tests and questionnaires were administered to the five classes in June 1971.

Name of test/questionnaire*	Time
1. TAT	40 mins.
2. Divergent Thinking Test	40 mins.
3. Moray House Intelligence Test	45 mins.
4. Non-academic Accomplishments Ques.	40 mins.

Information regarding father's occupation was obtained from school records. This was based on a survey done by the house-staff of the school about a year before testing for the present study was carried out. Other information obtained from school records was

1. English and Arithmetic marks from the last school examination in May-June 1971.
2. Teacher's comments and remarks about interest and participation in extra-curricular activities.

---

*A detailed description and rationale for each of these is given in the next chapter. Copies of tests and questionnaires are in Appendix B.

The Careers Office of the Local Education Authority collects information from pupils in the third year and onwards about their career aspiration and extra-curricular activites and interests. As some of this information was relevant for the present study specific information on the following questions was obtained from the Careers Office records:*

1. The job choice of pupils if this had been indicated either in interviews or in record forms. It will be noticed in the tables in Appendix A that information on this item was not available for all pupils as some of them had not decided what they wanted to do.

2. Information regarding leisure time activities and interest.

As the school has 35 or 40 minute periods, testing was done in double periods which these classes get for their English lessons. It can hardly be denied that most testing arrangements made for researchers in schools are rather arbitrary from the pupils' point of view in so far as they are usually informed, not long before the testing session that instead of the scheduled lesson they would be doing something else. It is not therefore unreasonable to assume that at least some pupils may be more reluctant participants than others in a particular testing session if the testing arrangements mean foregoing a favourite lesson.

---

*I am grateful to Mr. T.D. Black, Principal Careers Officer and his staff for their help and co-operation in enabling me to get this information.

As it happened, the decision to carry out testing in double English periods worked out rather well because the first two tests were the TAT and DT, both described as tests of "imagination" and therefore perhaps seen by the pupils as not very different from the type of work often done during English lessons. This also meant that there was less likelihood of pupils having to miss a very different kind of favourite lesson such as science or typing or cookery.

When I saw a class for the first time I introduced myself as a research student from the local University, gathering information on how pupils perform on certain tasks that are rather different from usual school-work. The pupils were also informed that whatever they wrote would be treated in confidence, and no one would be identified by name when the findings were reported.

To what extent these introductory remarks were instrumental in establishing the necessary rapport for the testing that was to follow is difficult to say. The importance of rapport is invariably stressed in texts on psychological testing (Anastasi 1968 Cronbach 19 60) and in test manuals, but the information on this subject is mostly indirect or incidental. For example, high reliabilities of most intelligence tests are often taken to indicate their immunity from unknown sources of performance fluctuation, or lack of rapport (Vernon 1969 p.110). Given the type of items of basic logic and reasoning which most intelligence tests contain, this is probably true. But measures like the TAT or the open-ended tests used in this study may be more susceptible to wide fluctuations in performance which are related to the degree of rapport effectively established or the lack of it in a particular testing situation.

For example, Hudson (1968) reports a quite unexpected increase in response output on an open-ended test as a consequence of what, according to conventional testing procedures, would be considered an extremely undesirable incident. This is how he describes the testing session from which "Some Fortuitous Evidence" emerged:

"I was in a bad temper and at the first sign of ill-discipline raised my voice more than I meant to, demanding that the miscreant pull his wits together on my behalf. There followed an hour and a half of what every teacher will recognise: uneasy, slightly resentful calm". (p.89).

The result according to Hudson was an exceptionally good performance by this group who were not significantly different from other groups he had tested under more normal testing conditions. Commenting on this unexpected finding he suggests that his "evil mood" may have "spurred the boys to try harder" or"It may also have helped to dissipate the air of uncertainty with which an open-ended task like Uses of Objects is surrounded: rather than dithering, unsure of what to write, they pitched in". It is relevant to ask if the results would have been the same had the boys been doing a more conventional type of test where no amount of "pitching in" would have earned them scores unless they knew the right answers. Possibly, in tasks of an open-ended nature, personality variables interact with situational cues to produce fluctuations in performance to a greater extent than they do in tasks of a predetermined type.

Following the publication of Wallach and Kogan's book, the studies by Kogan and Morgan (1969) and Kogan and Pankove (1972) certainly suggest that this question of interaction is extremely complex.*

---

*This digression on the question of rapport as being particularly problematic where open-ended tasks are involved seemed necessary in view of the fact that two of the tests used in the present study (TAT and DT) are of this type and some of the findings can be understood better in the light of this discussion.

CHAPTER VI

DESCRIPTION OF TESTS AND QUESTIONNAIRES

TAT

The first test to be given was Murray's (1943) Thematic Apperception Test, but instead of the full series of 20 pictures, only 6 were used. These were numbers 1, 2, 3BM, 7GF, 7BM and F13. In the test manual, Murray has provided the following description for his pictures.

Slide Number	Description
1	A young boy is contemplating the violin which rests on the table in front of him.
2	Country scene : in the background a man is working in the fields and an older woman is looking on.
3BM	On the floor against a couch is the huddled form of a boy with his head bowed on his right arm. Besides him on the floor is a revolver.
7GF	An older woman is sitting on a sofa close beside a girl, speaking or reading to her. The girl who holds a doll in her lap is looking away.
7BM	A grey-haired man is looking at a younger man who is sullenly staring into space.

The Sixth picture F13 shows a boy sitting at a table with an open book in front of him. In the background, near a half-open door can be seen two girls looking on.

The whole question of the value of TAT pictures is still largely unresolved, but it is now generally recognised that, given the complexity of the perceptual process, the objective properties of stimuli can no longer be considered "as determiners of perception" (Murstein 1963). An alternative way of determining the cue value of the stimulus is to define it in terms of the response it elicits. McClelland et al (1953) describe a study of this kind carried out by Veroff (1950) to determine the effect of achievement aroused on female high school students in response to pictures containing male and female figures. The general conclusion from this study was that female pictures have a low cue value for arousing achievement imagery, even for women. But as Murstein (1963) points out "defining the stimulus by its response necessitates the assumption that the response is not influenced by factors other than the "official" stimulus. If a subject does not evince hostility in a story told to a card, does it mean that the card contains no hostile characteristics? Is it not possible that the arousal of another motive than that depicted by the stimulus may determine the response?" (p.170). Indeed, this is what Field (1951) found in another study, also quoted by McClelland et al (1953), in which he managed to obtain higher nAch scores from women by referring to the variable of social acceptability in the "aroused" condition instead of using qualities of leadership and intelligence as had been done by McClelland and his co-workers. Commenting on this finding McClelland et al (1953) conclude that "the data unequivocally support the hypothesis that women's n Achievement is tied up with social acceptability, men's with leadership capacity and intelligence" (p.181).

A third, and now more widely used procedure for determining the stimulus value of pictures is to ask either expert judges or a large group of subjects representative of the population for whom the tests are to be used, to rate/rank pictures for the presence of particular "themes" or "concerns" chosen on a priori basis. Birney (1958), Haber and Alpert (1958) and Jacobs (1958) employed this method in their studies of cue characteristics. Using such indices as mean rank assigned to a number of pictures for the presence of certain "themes" or "concerns" (Achievement, Affiliation, Guilt etc.) and average intercorrelations of judges rankings of a picture on all the themes, these authors conclude that some of the pictures have a stronger stimulus "pull" than others, and therefore by implication, these high stimulus pull pictures are also less ambiguous in terms of the themes they suggest.

Determining the ambiguity of the TAT pictures and how this affects the type of story that is written to them, seems to have become a major concern of reasearch in this field. Murstein (1963) provides a comprehensive review of findings in regard to this question as well as about how thematic modifications, variations in lighting, focus, colour etc., influence story writing. The findings most relevant to the selection of pictures for the present study are summarised below.

To the question, "which is of greater utility : a test of great stimulus complexity capable of educing responses covering many different traits or a test scaled to tap only a single trait?"

Murstein's answer is that if "we want an overall measure of achievement, we might include scenes of home, school, job and social occasions in our measure. If it is desired to make a prediction about a specific situation (e.g. achievement motivation in a school setting), it would be helpful to portray scenes as close to the situation as possible. In sum, the consensus of opinion is that predictability is a function of the similarity of the test situation to the situation being predicted, a conclusion increasingly embraced by "objective" test constructors as well" (pp.178, 179-181).

In contrast to these meticulous methodological considerations for determining the cue value of TAT pictures, Arnold (1962) has argued that within quite a wide range of variation, it is immaterial what particular pictures are used to elicit a story. She herself has used pictures from Life Magazine for this purpose and reports finding little difference in the stories written to these pictures in comparison with those written to conventional TAT cards. Since Arnold uses a different scoring system from that used by McClelland et al (1953) it is possible that her claim may be valid for this particular scoring system only, especially as in her system credit may still be given for a protocol in which achievement imagery as such is absent and the dominant theme is either about attitudes to Right and Wrong or about Reaction to Adversity or about Human Relations. It is understandable that under such a scoring scheme the pictures need not necessarily carry achievement cues to elicit the appropriate imagery from subjects.

In view of these diverse conclusions regarding the stimulus properties of TAT pictures, the selection for the present study was made on the basis of rather rough and ready criteria which seemed to be of some significance within the general framework of this research. For example, since the subjects were 15 to 16 year old boys and girls it was considered reasonable to include cards which provided equal opportunity for identification with the same sex character (if indeed such identification does take place, the evidence is not clear at all on this question as Murstein's review shows pp.202-210). It was also felt that at least some pictures should depict scenes of achievement related activities with which a group of this type would be familiar, hence the selection of pictures 1, 2 and F13. At the same time pictures 3BM, 7GF and 7BM were chosen for their possible depiction of personal relationships, thus providing a balance, especially as Arnold's (1962) scoring categories were also to be used.

Administration of the TAT:

The test was given under group conditions, projecting each slide for 30 seconds and then allowing six minutes to write a story about it. In the terminology of McClelland et al (1953) the procedure used for administering the TAT may be described as "neutral", which is "neither to depress nor to increase the level of motivation but to keep it "normal", so as to obtain a measure of the motivation subjects brought with them to the situation. In other words to measure the motivation level elicited by the cues of everyday school life" (p.101).

The instructions read out to the pupils along with a sample page from the test booklet are reproduced in Appendix B. It will be noticed that the four questions around which the story to each slide is to be written, are repeated on the page represented here. As one page was considered enough for each story, these questions were repeated on all the six pages of the test booklet.

Scoring of the TAT:

Two rather different scoring systems were used in the present study. One was scoring system D-2 which is a modified form of McClelland's original scoring scheme. It has been described and used in several validation studies by Ricciuti and Clark (1957), Ricciuti and Shultz (1958), Saddaca, Ricciuti and Swanson (1956). It differs from the other more widely used system C in that there are no negative scoring categories used in this system and the category for scoring the presence or absence of achievement imagery (AI) has been expanded to specify the different broad headings covered by this category. Since the decision about whether or not achievement imagery is present in a story is most crucial for determining the rest of the score*, system D-2 which is more specific on this question was chosen as being more reliable.

The categories used for scoring system D-2 (Ricciuti and Clark 1957) along with the scoring sheet (Saddaca, Clark and Ricciuti 1957) are reproduced on the next four pages. From these it can be seen that a story may be scored for the presence of achievement imagery (AI) if it mentions:

---

*If it is decided that achievement imagery is not present then no further scoring is done and a score of 0 is assigned to the story.

1. Involvement or affective concern with achievement (INV).

2. Implicit or explicit evaluation of performance (E-1 to E-3).

3. Explicit positive or negative affect over mastery or failure (Ego Set + or -).

If on the basis of the above guidelines, it is decided to score a story for the presence of achievement imagery (AI), then a search is made in the story for the presence of any of the sub-categories under the heading "nAch Scoring Components" on the scoring sheet. The final score for a story is the sum of the sub-category scores, with the exception of sub-categories I +,?, -.

Saddaca, Ricciuti and Swanson (1956) report product-moment correlations of .74, .81 and .83 among three judges for total scores derived in this way. Other workers report inter-scorer reliability (in terms of product-moment r) of .64 to .95 using one or other of McClelland's scoring systems (Murstein 1963).

Scoring System D-2 Categories Used in Analysing
Stories for Achievement-Motivation Imagery

*INV	Involvement of affective concern: scored whenever any of the categories N+ to I_4 are scored.
*E-1	Explicit evaluation of performance ("he did a good job").
*E-2	Implicit evaluation of performance -- unique accomplishment ("he creates a masterpiece").
*E-3	Implicit evaluation -- involvement in activities affecting self-esteem ("he studied hard").
*Ego Sat +	Explicit positive affect over mastery ("he's proud of his grade").

*Ego Sat - Explicit negative affect over failure to achieve mastery ("he felt disheartened over his school failures").

I+, I?, I- Used to indicate whether net outcome of instrumental activity is positive (successful), indeterminate, or negative (failure).

AI Achievement imagery (AI) (first judgement made).

N+ Need to achieve ("he wants to be an engineer"; "he hopes to win").

*N- Need to avoid failure ("he does not want to fail the critical test").

Ga+ Anticipation of future success ("he knows he will solve the problem").

*Ga? Concern about future ("he wonders whether he'll make the grade").

Ga- Anticipation of future failure ("he's worried about failing the course").

G+ Positive affect over success ("he is happy about winning the prize").

G- Negative affect over failure ("he is disgusted with himself for failing").

I Instrumental activity, i.e. activity directed toward an achievement goal. Sub-categorized as follows: *I_1 "working carefully"; *I_2: "working hard"; *I_3: "working for a long time";*I4: "working fast"; *I_u: "unmodified" instrumental activity, i.e. just working."

Nup Nurturant press ("the foreman helped him to learn the job").

*PA/

*PA	Personal asset ("he was an alert and conscientious student").
B	Blocks or obstacles which interfere with or slow down goal-directed activity. Sub-categorized as: Bp: personal block ("he failed because of his laziness"); Bw: world block ("the storm kept him from getting through on time"); *B?: not identifiable as either Bp or Bw.
AT	Achievement thema: Scored if story primarily deals with an achievement theme, provided two of the categories from N+ to I_4 are present.

---

*Indicates those categories absent from scoring system C.

Note: Categories INV to I+, I?, I- do not presently contribute to achievement motivation scores. (cont. on page 136).

Date________ Scorer________ Sample Designation________ Subject Name________ Subject No.________

not scorable	AI Determinants						Picture No.	nAch Scoring Components																							Picture No.	Remarks
	involvement	Evaluation			Ego Satis			AI	Involvement															Aids Blocks					Ach Th	TOTAL		
									N		Ga			G		I																
		1	2	3	+	-			+	-	+	?	-	+	-	U	1	2	3	4	+	?	-	Nur	PA	Bp	B?	Bw				
							1																								1	
							2																								2	
							3																								3	
							4																								4	
							5																								5	
							6																								6	
							7																								7	
							8																								8	
							9																								9	
							10																								10	
							11																								11	
							12																								12	
							Tot																									

The second scoring system used was that of Story Sequence Analysis, devised by Arnold (1962). This system is also considered by its author as "A New Method of Measuring Motivation and Predicting Achievement"*, but under this system there is provision for scoring every story, whether achievement imagery is present in it or not. In fact it is not the story as such that is scored, but its "import" which is first extracted from the story in such a way that it conveys the central idea or the basic attitude contained in the story. If the "import" of a story is not clear on its own, the general trend of thought in the stories before and after it may be used to derive the import of that story. Once all the imports have been extracted these can be scored on a five point scale from -2 to +2 (with zero as the mid-point under one of the following four headings:

I - Achievement, Success, Happiness, Active Effort (or lack of it).

II - Right and Wrong.

III - Human Relationships.

IV - Reaction to Adversity.

---

*This is the subtitle of the book.

Each of these categories is further divided into sub-categories*. For example, an import concerned with Achievement (I) may be scored +2 under sub-category I.A.1. "Success is reached when goals are reasonable...". If the import is only mildly suggestive of achievement oriented goals it may be scored +1 under sub-category I.A.1. "Goals are minor or achievement is yet uncertain." Similarly, if an import is strongly negative it may be scored -2 under I.A.1. "Success follows upon action for negative motives." Thus the fact that the mere presence of achievement imagery is not enough for giving a positive score is one way in which Arnold's method may be considered an improvement on the more traditional scoring systems in which a positive score is given even for the presence of such categories as "Concern about the future" (Ga?) and "Anticipation of future failure" (Ga-).

The total score for an individual is the algebraic sum of scores obtained on each story. To make this score comparable when derived from protocols of different lengths (in terms of the number of stories used) it is converted into a Motivation Index which is a ratio of the

---

*The main categories and headings along with the sub-categories for one of the headings are reproduced from Arnold (1962) at the end of this section. (pp.143-148).

obtained score units* divided by the total possible obtainable score units on protocols of a particular length. Unless a protocol is awarded the maximum possible score for its length, the ratio is usually in the form of a decimal figure. For ease of computation Arnold multiplies this ratio with a constant value of 200. Thus, regardless of the number of pictures used, the Motivation Index ranges from 0 to 200, with 100 as the turning point; scores below 100 indicating negative motivation, those above 100 indicating positive motivation. For example, on a six story protocol, as in the present study a score of -4, +8 = +4 would cover 16 score units since the total range would extend from -12 to +12 = 24 score units. Dividing the obtained score units by the total obtainable units we get $\frac{16}{24} = .666$, this multiplied by 200 gives a Motivation Index of 133.

Arnold (1962) provides a conversion table for scores ranging from -40 to +40 and for protocol lengths of 10, 11, 12, 13, 14, 15, 16 and 20 stories. For the present, study due to the necessity of keeping testing time to a minimum so as not to take pupils out of their regular lessons more than absolutely necessary, it was possible to use only six TAT pictures. However, this may not be considered a serious omission. Reitman and Atkinson (1958) report that adding pictures beyond a certain number may actually lead to lower reliability. The standard instructions given to subjects describe the TAT as a test of imagination.

---

*Arnold does not make this distinction very clear, but it seems that a score unit is a point on the continuum of the score range from the negative to the positive end and not just the sum of the negative and positive scores. For example on a 20 story protocol where the possible range is of 80 score units extending from -40 to +40, a score of +20 would cover 60 score units that is, 40 units on the negative side and 20 on the positive.

As Murstein (1963) points out "after a given number of cards, the subject may seek to vary his themes so as to avoid telling the same story. The effect of this manoeuvre would be to lower split-half reliability" (p.137). Ricciuti and Clark (1957) also found that the effect of their "achievement oriented" instructions petered out after the eighth picture in a 12-picture series.

Since Arnold's method of Story Sequence Analysis was to be used in the present study, the number of pictures was raised from the usual four to six, although it still falls short of the minimum ten recommended by Arnold (1962). Arnold's conversion table does not provide values for less than ten stories, so to obtain the Motivation Index for this study the total score was doubled and its equivalent Motivation Index for a twelve story protocol (thus also doubling the number of pictures used) was read. In other words, it was assumed that the six stories actually written were a reasonably good sample of the type of stories that would have been written had there been time to obtain more than these six.

Arnold gives reliability estimates of this scoring method in terms of the percentage of inter-scorer agreement. They range from 94% to 97% among three judges. As this is a comparatively new method which has not been used by many workers other than Arnold's own research team, intra-scorer and inter-scorer reliability checks were made for the present study. After a 10-month interval I scored again a random sample of 20 scripts (120 stories) and the rank order correlation between the first and second set of scores was $r = .628$ ($p < .005$ one-tailed).

Inter-scorer reliability between an experienced scorer and myself was r = .975 ($p<.05$ one-tail) and r = .586 ($p<.005$, one tail) between another scorer* (who learnt the scoring system for this study) and myself. It looks as if for a projective test of this kind a reasonable level of scorer reliability can be established provided the markers are well acquainted with the scoring system.

There were three main reasons for using a second scoring system in the present study:

(1) Although McClelland et al (1953) do not report any direct investigation of the relationship between nAch score and verbal fluency (in terms of length of story) later research (Child, Storm and Veroff 1958, Ricciuti and Clark 1957 and Walker and Atkinson et al 1958) has consistently shown a significant positive relationship between nAch score and length of protocol. As Arnold (1962) points out, given McClelland's scoring systems, this is inevitable, since once a decision has been made about the presence of Achievement Imagery (AI) in a story, the rest of the score depends on the number of sub-categories scored. The longer a story, the more likelihood there is of the presence of a greater variety of sub-categories. As the pupils in the present study came from a wide range of ability, it was considered possible that their lack of verbal fluency might affect their nAch scores and therefore Arnold's scoring system was also used as a check on this point.

---

*I am grateful to Lawrie Maloney, a fellow research student for spending a considerable amount of time in familiarising himself with the system and scoring the protocols for me.

As Table 1.1 shows, scores derived from Story Sequence Analysis (MI) vary from class to class independently of IQ which is a measure of verbal reasoning* ability, whereas the scores based on D-2 system of scoring (nAch) follow the pattern of variation in IQ from class to class. This trend is further confirmed in the overall correlational analysis in Table 2.1 in which MI has no significant relationship with IQ ($r = .078$,ns) but nAch is positively related to IQ ($r = .205$, $p < .001$). It should however be noted, as Ricciuti and Clark (1957) point out, that a positive relationship between verbal fluency and achievement motivation need not necessarily be considered a "contaminating" factor. On the contrary it may indeed be of some significance as a correlate of achievement motivation. This alternative is discussed in some detail in Chapter VII where a cognitive theory of motivation is considered.

(2) The second reason for using Arnold's method of analysing TAT stories was that scoring system D-2 by not scoring those protocols which did not have Achievement Imagery (AI), seemed to be rather wasteful of that may well have been quite rich psychological data. It will be recalled that one of the aims of the present research was to isolate four experimental groups (High-High, High-IQ, High-DT, Low-Low) and see if there are any systematic differences among them in terms of their attainments, interests and motivation.

---

*The distinction between verbal reasoning and verbal fluency is not overlooked here, but in view of the strongly verbal bias of the intelligence test it seems reasonable to assume that the two are closely related.

As mentioned before, Arnold's method provides for categories other than that of achievement only. In view of the research findings of Barron (1969), Getzels and Jackson (1962), Hudson (1966, 1968) and Wallach and Kogan (1964) regarding the personality correlates of divergent thinking ability, it was expected that the four experimental groups would differ from each other on the remaining three of Arnold's scoring categories, namely Right and Wrong, Human Relationships and Reaction to Adversity. More specifically, it was hypothesized that divergent thinking would be associated with (a) more flexible attitudes to Right and Wrong, (b) more open expression of feelings in Human Relationships and (c) more emotional and a less stoic Reaction to Adversity.

(3) The third and final reason for using the two scoring systems was to see how they related to each other. Arnold herself does not provide any data on this, although she argues that her system is likely to have better predictive validity than McClelland's (Arnold 1962, pp.17, 24).

ARNOLD'S SCORING SYSTEM : Categories and Headings

I. ACHIEVEMENT, SUCCESS, HAPPINESS, ACTIVE EFFORT (OR LACK OF IT)

A. Goals, purposes.

B. Means taken toward goal.

C. Adaptability as to goals and means.

D. Influence of others on success, achievement, etc.

E. Consequences of success (failure).

F. Attitudes connected with success (failure).

II. RIGHT AND WRONG (Well-intentioned, reasonable, responsible action versus ill-intentioned, impulsive, harmful, irresponsible action)

A. Actions.

B. Intentions, attitudes, emotions.

C. Effects (consequences) of punishment.

III. HUMAN RELATIONSHIPS

A. Good (friendly) relations (including friendship, love, marriage).

B. Bad relations (including quarrels, enmity, etc.).

C. Influence of others.

D. Influence of others on success, achievement, etc.

E. Influence on others.

F. Attitudes (toward people and things, God, Nature, Life, etc.).

G. Attitudes connected with success, achievement or lack of it.

IV. REACTION TO ADVERSITY

A. Loss, harm, danger, terror, separation, disappointment, difficulties.

SCORING SYSTEM : Individual Scores

I. ACHIEVEMENT, SUCCESS, HAPPINESS, ACTIVE EFFORT (OR LACK OF IT)

A. Goals, purposes

+2

1. Success is reached when goals are reasonable; it follows upon
   a. action for ethical, religious, well-intentioned motives
   b. action dictated by prudence, experience, etc.
2. Failure, no achievement, when goals are unreasonable or self-centred; it follows upon
   a. action for ill-intentioned, imprudent motives
   b. failure to act for ethical, religious, well-intentioned, prudent motives
   c. action undertaken to impress others

When the import indicates neither success nor failure, look for evidence of:

3. Preference for immaterial values, as against material, expedient, irrational values; preferences for values that are
   a. ethical b. religious c. spiritual d. altruistic
4. Optimism, implying
   a. constructive action (e.g. life is responsible, constructive, worthwhile; compromise on principles leads to disaster, harm, penalty, etc.
5. Imports exemplifying an active personal relation to God
   a. God is seen as creator, father, sustainer of life
   b. readiness is expressed to do His will

I. ACHIEVEMENT, SUCCESS ... A. Goals, +1, -1

+1

1. Goals are minor, or achievement is yet uncertain:
   a. success when goals are modest (e.g. you may not become famous but you'll do well)
   b. success with some failure along the way
   c. goals striven for but outcome not certain

When the import indicates neither success nor failure, score for:

2. Imports embodying (constructive) principles (e.g. freedom must not be sacrificed for strength)

3. Optimistic imports:
   a. with reasons given, but not implying action (e.g. in nature, good times follow after hard times)
   b. implying that pessimism is undesirable (e.g. a dim outlook makes things seem worse)

4. Imports appreciating immaterial values (e.g. education, learning, etc., is valuable)

-1

1. Lesser goals are preferable; because they
   a. require less effort b. do not affect personal worth c. are best
2. Two conflicting goals can be reached:
   a. with the help of others b. by chance, fate
3. Success follows action for extraneous motives:
   a. for the approval of others
      i. simply to please others
      ii. to please others by delaying one's action
   b. for fame or recognition
   c. for the sake of conformity
   d. for self-centered motives (e.g. you succeed if you look after your own interests)

4./

4. Success is foretold if character should try, persevere, etc.

When no indication of success or failure, score for evidence of:

5. Optimism without good reason ("Pollyanna" stories)

   a. success comes as eternal reward (no action)

6. Heroics, phoniness of every kind

-2

1. Success follows upon action for negative motives

   a. involving ill-intentioned or self-seeking goals

   b. failing to act for ethical or well-intentioned motives

   c. delaying when immediate action is called for

   d. acting to impress others (e.g. showing off)

2. Success is possible:

   a. is uncertain, a mirage

   b. is hoped for oneself or others

   c. is dreamed about or thought about

   d. comes in unexpected guises (e.g. you dream of one thing, become another)

   e. is expected but failure is experienced instead

3. Success if foretold from the manner or look of the character (e.g. I can tell he'll be successful from his determined look)

4. Failure as outcome; failure is:

   a. expected

   b. experienced, just happens, etc.

   c. not admitted (e.g. everything will turn out well - when story indicates (failure)

   d. caused by other people or things

   e. result of chance, fate, etc.

When the import indicates neither success nor failure, score for evidence that

5. Goal is not firmly pursued:

   a. it seems foolish, unrealistic

   b. is relinquished because of pain, danger, etc.

   c. becomes more difficult to reach

   d. is wondered about

6./

6. No goal is indicated

7. Pessimistic imports (e.g. when destruction is general, you may just be able to save yourself, but there is no help even for the man next to you).

<u>Divergent Thinking Test</u>

Description of Tests:

This was the second test to be given and consisted of three sub-tests, namely Uses, Similarities and Consequences. All three are derived from Guilford's (1967a) research on divergent thinking abilities as part of his structure of intellect model, but the actual format of the sub-tests is taken from different sources. The objects in the Uses sub-test are the same as those used by Getzels and Jackson (1962).

In this sub-test the subjects are given the names of a few common objects such as a brick, a pencil, a toothpick etc., and asked to write the uses that they could think of for these objects. The "items" for the Similarity sub-test are taken from Wallach and Kogan (1965). Although this test in its convergent (one right answer) form has been part of conventional intelligence tests (WAIS 1955, W-B II 1946, and Stanford-Binet Intelligence Scale 1962) Guilford (1967a) and Wallach and Kogan (1965) have used it as an open ended test. The difference in its use as an open-ended test lies in the scoring procedure whereby a person can gain more than one point for mentioning the different ways in which two things are similar. Although in its traditional forms more than one basis of similarity is recognised in the answer key, no extra credit is given for mentioning more than one of these. For the third sub-test, Consequences, the "items" were written specially for this study. The idea behind the sub-test is to see how far subjects can stretch their thinking to foresee the changes which would follow if certain hypothetical situations were to arise.

Administration of Divergent Thinking Tests:

The format in which the test was presented to pupils is given in Appendix B. As the instructions at the beginning of the test and for the three sub-tests show, an effort was made to encourage freedom of expression in doing the test, although due to administrative constraints* the test had to be given under timed and group conditions. It will also be noticed that there are altogether only fourteeen "items" in the test, which, given the 40 minutes for doing it, works out at about 3 minutes per "item".

---

*i.e. the necessity of completing the testing within an alloted number of time-table periods.

Kogan and Morgan (1969) and Leith (1972) report giving the same amount of time (3 minutes) per item when they administred their tests under timed conditions. Hudson (1966) also reports that given five objects in the Uses test, the majority of his subjects had finished within 15 minutes. At the end of each 40 minute testing session I also made a rough visual check round the room and found that most pupils had stopped writing, although the instructions clearly said that they could go back to an "item" if they so wished. On the whole, my impression is that the time limit and the group administration of this test did not seriously affect the pupils' output.

This impression is further strengthened by the findings of Kogan and Pankove (1972) in a recent study in which individual and group administration of open-ended tests to a follow-up sample of Pankove and Kogan's (1968) initial group showed that individual administration lead to a significant drop in the boys' scores. In another study with nine to thirteen year olds Leith (1972) also found that of three open-ended tests given under "No Stress" and "Mild Stress" conditions, one (Word Association) showed significantly higher mean scores under the "Mild Stress" condition and there were no significant differences in the other scores under the two conditions. Vernon (1971) has argued that valid scores on open-ended tests can be obtained under group conditions if the task is defined clearly, and "if the tester is not a teacher nor identified with the school" (p.256). But the issue is not a simple one, as the findings of Elkind, Deblinger and Adler (1970) have shown. According to these authors even under the same conditions of administration, children who had been taken away for testing while engaged in an "interesting" activity had lower mean scores in comparison with those who had been interrupted in the middle of an "uninteresting" activity.

The implication of this study is that it is not only test administration as such, but its antecedent conditions too which affect performance.

Three main points emerge from these studies:

(1) Kogan and Pankove's (1972) study suggests that age may be an important factor in determining the effect of a particular testing context on divergent thinking test scores. That is, at the primary/elementary school level a face to face individual testing procedure may be conducive to better performance but for adolescents this may be an inhibiting influence. Vernon (1971) also concludes that "test scores obtained under relaxed conditions show their greatest superiority in relation to age (i.e. young pupils did better in the relaxed than in the formal group)". The present study did not involve pupils from a wide age range; the mean age of the third and fourth year pupils was 15 years, 1.5 months and 16 years, 0.3 months respectively. As Table 1.3 in Appendix A shows, there is no significant difference in the divergent thinking scores of the fifteen and sixteen year age-groups. It therefore looks as if age differences have to be considerably more than in the present study to make any significant difference to performance on divergent thinking tests. It may also be that with pre-adolescent subjects as in Leith's and Vernon's study even smaller age differences affect performance, but at the fifteen and sixteen year level the effect of age on performance begins to level off.

(2) A "relaxed" atmosphere during testing does not necessarily require an untimed and individual testing situation. Even under group testing a "relaxed" atmosphere may be maintained if (a) sufficient time allowance is given, (b) the test administrator is not identified with the authority figures in school (Vernon 1971) and (c) if the instructions emphasise evaluation of the test rather than of the subjects as the object of testing (Leith 1972).

(3) As Kogan and Morgan's (1969) and Leith's (1972) findings suggest, it is not even certain that a "relaxed" testing context is always the most suitable one for obtaining valid divergent thinking scores. In the study referred to above Vernon also found that "the factorial structure of formal and relaxed scores is generally similar" in the two testing conditions.

Writing about the role of motivational factors in creative accomplishments Maddi (1965) has also argued that in the literature on creativity, an old wives' tale which has survived for far too long is that a completely non-evaluative and supportive environment from which all sources of anxiety and frustration are absent is necessary for creativity to thrive. Maddi suggests other characteristics such as a preference for novelty and attending to the "internal environment" of feelings and thoughts as the more likely sources of creative performance. Hudson (1966) goes further and thinks that although "progressive schools do make most children happier that authoritarian ones; they withdraw from children the cutting edge that insecurity, competition and resentment supply" (p.134).

At an empirical level, his own evidence (referred to in Chapter V) regarding the exceptionally good performance of a group with whom he had lost his temper, supports the proposition that a relaxed testing context does not always elicit the best performance on open-ended tasks.

It seems therefore that Wallach and Kogan's (1965) conclusion that an individual, non-evaluative and untimed testing context is crucial for obtaining valid scores on divergent thinking tests is open to doubt as a generally applicable proposition.

Scoring of the Divergent Thinking Tests:

All sub-tests were scored for the number of <u>different</u> uses, similarities and consequences mentioned and for the number of <u>unique</u> responses given. In other words, the two scoring criteria were flexibility and originality. Both these criteria emphasise the quality of ideas given rather than sheer quantity. Although scoring for quality involves a much more time consuming procedure this method was used in preference to a simple fluency score because on a priori grounds it seems more valid, in the sense that it favours the person who has many different ideas rather than the one who goes on writing about the same idea in a repetitious way. For example, when asked to give different uses for a brick if one person went on giving many answers centred round the idea of building (houses, walls, pavements, steps, roads etc.) and another gave a varied list of uses (door stop; weight - to drown or hold things down; goal-post or marker; build words; stepping stones in a river etc.) the latter set of answers seems qualitatively better as it indicates a more resourceful and imaginative approach.

In the description of his structure of intellect model, Guilford (1959a) also considers flexibility to be the most crucial aspect of divergent thinking ability: "In divergent thinking operations we think in different directions, sometimes searching sometimes seeking variety.....".

At the empirical level too Guilford and his co-workers (Frick et al 1959, Wilson et al 1954) report finding a factor of "spontaneous flexibility" which reflects the number of different categories of classification for an item. There have been few direct studies of the relationship between sheer fluency in terms of total number of responses and flexibility in terms of <u>different</u> responses. A recent study by Hargreaves and Bolton (1972) is one in which it was concluded that "Fluency, Flexibility, Originality and Elaboration subscores [on the Minnesota Tests of Creative Thinking] were highly inter-correlated" (p.451). The correlations between Fluency and Flexibility ranged from .63 to .95, all significant at .001. Some incidental information on this question is also available in Cicirelli (1965) who found that factorially verbal fluency-flexibility-originality belong together, and in Kogan and Morgan (1969) who report correlation ranging from .47 to .98 (all significant at .01) between the fluency and flexibility scores in their study. In the light of the quite close relationship between fluency and flexibility shown in these studies it seemed reasonable to assume that using flexibility scores for the present study would not entail any serious loss of information on the divergent thinking ability of pupils.

In his review of the creativity literature, Wallach (1970) has repeatedly argued that, in Guilford's terminology, "ideational fluency" rather than "spontaneous flexibility" is the ability which "may be paradigmatic for the kind of cognitive performance that is maximally cohesive in itself and maximally distinguishable from convergent thinking" (p.1221). Although Wallach discusses a number of studies to support his argument the main line of reasoning throughout seems to be as follows:

(1) Studies which have scored open-ended tests for ideational fluency have found low correlations between IQ and divergent thinking and comparatively higher intercorrelations among the open-ended tests themselves;

(2) Guilford's theory (1950, 1956, 1959, 1967) on which most of the present creativity research is based, postulates precisely such a pattern of relationships between convergent and divergent thinking abilities; therefore,

(3) A scoring scheme which produces the predicted distinction and associations must necessarily be the most valid one for obtaining an index of divergent thinking ability.

There seems to be a circularity in this argument if we look closely at the question of the validity of ideational fluency scores. The very independence from IQ which Wallach considers as supportive evidence of divergent thinking ability may also imply a lack of any qualitative check on the type of responses given by someone high on ideational fluency. For example, the ability for evaluation and self-criticism is generally considered an important aspect of intelligence and creativity, and it can be argued that a person with

a/

a high ideational fluency score may be lacking in this ability if he goes on giving repetitious answers to open-ended questions without any set-shifting or evaluation*. The usual criterion for ideational fluency scoring is that the response should be "relevant" or "appropriate" and not be "bizarre". On its own a response may meet all these requirements, but considered as part of a string of ideas it may not be deserving of a credit. For example, the Uses for a brick in the present study have often been given as: "to build houses, walls, offices, schools...." by the same person. Individually, each response is relevant and appropriate but ought this to get a score of 4, especially as the instructions had clearly called for as many <u>different</u> uses as possible?

---

*Heim (1970) makes a similar point in her discussion of the role of self-criticism and judgement in creativity, but she introduces the variable of social desirability also as one of the many criteria for judging the appropriateness of a response to open-ended tests: "If we consider the Uses for Things Test, for instance, it is likely that a fair number of the Subjects who think of "breaking a window" as a use for a brick, censor this. Such a subject will do less well on the test than will the uninhibited answerer who does not hesitate - but the former subject who both had the idea and decided to drop it may well prone to be the better Scientist" (p.41).

A person giving such a response would no doubt get a high score on a task where the number of responses given is the yardstick, regardless of their quality. Consequently, this scoring procedure would yield low correlations with intelligence, whereas a procedure which takes into account the quality of responses, as scoring for flexibility does, would have comparatively higher correlations with intelligence.

To illustrate the sort of difference that is involved in these scoring procedures, two sets of responses to the question:"What would the consequences be if human beings could fly as the birds do?" are quoted here from the present study.

Pupil A (IQ 103): There would be no - planes, aeroplane engines, airports, pilots, navigators, stewards, air-hostesses, runways, radars, landing-lights, control tower, fire/ambulance and airport police, customs, fares, adverts on T.V.

Pupil B (IQ 119): Could get from A to B quicker, build more tree houses, go for day trips to the moon, save the expense of cars/lorries/buses, catch birds easily, have wings, won't have to wear shoes so often, won't get feet wet from puddles, Walter Raleigh wouldn't have put his cloak down for Queen Elizabeth I.

In terms of the sheer number of responses Pupil A would get a score of 16, Pupil B,9.

A number of similar examples could be quoted where pupils of lower IQ have given their ideas in an associational string, without paying attention to quality, whereas those with comparatively higher IQ have given a more varied and ingeneous set of ideas. Obviously, if IQ/DT correlations are based on fluency scores of this kind the likelihood of obtaining low correlations is much more than if open-ended tests are scored for quality. It is therefore questionable whether the use of ideational fluency scores is justifiable simply because they yield low correlations between IQ and divergent thinking. Of course, face validity alone of tests and scoring procedures cannot be considered a sufficient basis for claiming a distinction, or lack of it, between convergent and divergent thinking; but as Wallach himself points out, it must be "taken care of" before empirical evidence is marshalled. It is with reference to face validity that a distinction is made here between fluency and flexibility scores and it seems to be a crucial distinction for determining the pattern of IQ/DT relationships which is itself still a central issue in creativity research.

Discussing his tests for ideational fluency, Guilford (1967) has clearly stated: "that a fluency score in all these tests is merely a count of the total number of relevant responses, indicates that DMU* is measured by sheer quantity of relevant output in a limited time... the quality of output, or more accurately stated, the quantity of high-quality responses is an indication of flexibility in production" (pp.142-143).

---

*Abbreviation for the cell "Divergent Semantic Units" in the structure of intellect model, M standing for Semantic (Meaning). Guilford has to do this to avoid using S again as he has already utilized it for the "Symbolic" cells.

In view of this statement, it is surprising that although elsewhere (Wallach and Kogan 1965) Wallach has argued that scores on open-ended tests given under timed conditions are not valid measures of creativity, in this later review (Wallach 1970) he considers ideational fluency tests as described by Guilford above, as being "paradigmatic" for obtaining the convergent-divergent distinction. One suspects that now Wallach has moved to a position where any procedure which yields the desired convergent and discriminant validity, statistically (to use Campbell and Fiske's terminology), is considered the most valid one.

The criterion of uniqueness used in scoring was that a particular response should occur only once in all the 108 scripts scored. This extreme definition of uniqueness was used instead of the procedure of weighting uniqueness scores on the basis of percentage of occurrence, because it is more straightforward and some workers in the field, (e.g. Wallach and Kogan 1965) have found that the cumbersome procedures used for weighting do not make any substantial difference to the ranking of subjects when compared with this simpler method. As Wallach and Wing (1969) point out, "The definition of uniqueness is a partially arbitrary matter" since a response that is unique with reference to one sample may not be so far another sample. It was due to this reason that for the present study a criterion of uniqueness was used which is specific to the sample of the study rather than based on "norms" derived from another sample.

To meet the requirements of the scoring scheme described above, it was necessary to tally every single response that was given and then (a)/

(a) on the basis of the frequency count determine which of these may be considered unique and (b) by a content analysis of the answers decide which ones may be "collapsed" together to form a single category for the flexibility score. When a list of common and unique responses had been prepared in this way, another judge* was asked to go through the list for all the "items" in the Uses sub-test and put a question mark next to the response which he considered had been:

(a) Unjustifiably excluded from a "collapsed" category, i.e. it was not very different from the others in that category. For example in the original list "Cut out design, origami patterns, doyleys" had been grouped separately from "Make Christmas decorations, paper chains, hats", but the uses suggested in both these instances for a sheet of paper were considered on review of the list to belong together.

(b) Unjustifiably included in a "collapsed" category. For example, for suggested uses of a brick, "Build residences, shops, offices, walls, paths, playgrounds", had all been grouped together, but the last two were considered by the second judge to be different from the rest as they referred to making of flat surfaces rather than buildings.

(c) Listed as unique because it had occurred only once, but he considered it inappropriate or bizarre.

---

*I am grateful to Mark Austin, a fellow research student for doing this for me.

Responses with a question mark were counted as instances of disagreement. The percentage of agreement worked out in this way was as high (94% for "sheet of paper", 98% for "brick" and 100% for "pencil", "paperclips" and "toothpick") as that reported by Wallach and Kogan (1965) and Kogan and Morgan (1969) for similar tests.

It is usual in studies of this kind to resolve disagreement between judges by mutual discussion before actually scoring the scripts. This seems a rather dubious post hoc procedure which inevitably yields spuriously high inter-judge correlations. In view of the close agreement between the judges on the scoring categories in this study, the second judge re-scored a random sample of thirty-one scripts for flexibility and uniqueness, using his own modified categories. The rank-order inter-judge correlation was .842 ($p < .0005$, one-tailed) for the total of flexibility and uniqueness scores. It was not possible to obtain the test-retest reliability for a sample of this size but a group of twelve pupils was available for re-testing on another occasion, eleven months after the main testing session. Rank-order correlation between the total of flexibility and uniqueness scores on the initial and subsequent testing was .730 ($p < .005$, one-tailed).

To get some estimate of the internal consistency of the divergent thinking test, intercorrelations among "items", sub-tests and totals were obtained. These are given in Table 2.3, Appendix A. It will be noticed that with the exception of column 5 and row 5 (Sheet of paper) where the correlations are exceptionally low, the other correlations fall typically between .3 and .6.

It will be recalled that the maximum disagreement between two judges on the scoring categories was precisely on this "item" and the low inter-correlations may well be due to unreliability in scoring. However, inspite of the fact that these low correlations must have pulled down the average, the over-all average r (using Fisher's z transformation) between individual "items" and the divergence total for flexibility and the divergence total for uniqueness were .62 and .52 respectively. In comparison with these average correlations, those between divergent items and IQ, English and Arithmetic were considerably lower, i.e. .30, .37 and .27 respectively. Considering that the item-total correlations are independent in the sense that scores of the particular item being correlated with the total were removed from the total before these correlations were computed, all correlations in Table 2.3 are comparable.

Raw score distributions for both flexibility and uniqueness showed considerable skewness as the following table shows:

Table VI.1 : Raw Score Statistics of Flexibility and Uniqueness Scores of Divergent Thinking Test*

	Mean	Mode	Stan. Dev	Range	Skewness
Flexibility	44.030	41.000	16.220	14 to 106	1.330
Uniqueness	5.545	2.000	4.897	0 to 26	1.524

It was therefore decided to rescale them to the mean and standard deviation of the IQ which had an almost normal distribution.

---

*Raw score mean and standard deviations for each item and for the sub-tests are given in Table 2.3 , Appendix A.

Moray House Intelligence Test:

This was the third test to be given. The actual test used was Moray House Intelligence Test (Adult)I revised in 1970. Like other tests of this series, it consists of multiple-choice and completion items of the objective type. Under the general title of "intelligence", the test contains items of classification, analogy, vocabulary, ordering, symbolic reasoning and verbal comprehension.

The administration and scoring was done according to the standard procedure given in the manual.

The test manual reports reliability coefficients of .945 and .833 for this test.

Mean IQs and standard deviations are given in Appendix A, Table 1.1 for individual classes as well as for the whole group.

Non-Academic Accomplishments Questionnaire:

Format of the Questionnaire:

This was a considerably modified version of the questionnaire used by Wallach and Wing (1969). The areas of non-academic accomplishments covered were the same eight used in the American Study - Art and Craft, Drama, Speech and Debate, Creative Writing, Music, Science, Social Service and Leadership with Sports and Games as an additional heading. At the end of the questionnaire, space was provided for mentioning any other achievement not covered by the rest of the questionnaire.

The wording and format of the questionnaire were different from Wallach and Wing's in the following way:

(a) The vocabulary and activities mentioned were changed to suit the local secondary school pupils, whereas the original research had been carried out on American under-graduates.

(b) Instead of a variable number of activities in each area, six levels of participation, in an ascending order of involvement, were used for the present study. Thus, the first level (a) usually referred to membership of a club or creation of an original piece of work, the next level (b) referred to recognition of the work within the school in a display or exhibition, the third level (c) referred to participation in a similar exhibition in a wider context than that of the school (i.e. local or regional level), next (d) came the winning of a prize or award at this level. The highest two levels were (e) selection for exhibition at the national level and then (f) gaining some form of recognition at this level. If the activity was of a type where actual participation rather than display of created work would be more relevant, the degree and level of participation was determined by the frequency of participation (sometimes, regularly etc.) as well as by the geographical area within which recognition in the form of prizes and medals had been awarded. The questionnaire is reproduced in Appendix B.

Administration of the NAAQ:

Like the tests in the present study, this questionnaire was also given to groups of pupils formed on the basis of school classes. If after the instructions had been read, any doubts still remained about what was required for completing this questionnaire, the points were further clarified.

After this, the pupils went through the questionnaire on their own and marked the statements that seemed to apply to them.

One school period (of 40 minutes duration) was sufficient for most of the pupils to go through the questionnaire. Those few (two at the most in each class) who had not managed to finish it were allowed to stay on and finish it.

Scoring of the NAAQ:

As mentioned before, every field of activity was represented by six statements, arranged in an ascending order in terms of the degree and level of participation or accomplishment in that field. The first statement in each area which referred merely to membership of a club or pursuit of a creative activity was given a score of 1, the next one 2 and so on until the final statement which represented the highest level of accomplishment was given a score of six. The rationale for this scoring scheme was based on the recognition that the quality of accomplishment in each field ought to be evaluated and therefore the statements representing a greater degree of participation and accomplishment were given increasingly greater weight. The total score of an individual was the sum of these weighted scores for each of the statements marked as being descriptive of that individual's achievements.

Although the scoring system used was quite objective it was still necessary to determine the test-retest reliability of this questionnaire as it has not been used before in its present form. Richards et al (1967) report K-R20 reliabilities ranging from .61 to .84 for a similar questionnaire.

However, Wallach and Wing (1969) from whose questionnaire the present one is directly derived, provide no information on its reliability. Neither do Cropley (1972) or Kogan and Pankove (1972) who have also used it.

There was no reason to believe that with a questionnaire of this type, where both the level and the area of achievement are not very clearly defined, the same level of reliability as with the original questionnaire (Richards et al 1967) would be obtained. On the other hand, unless sufficient confidence could be placed in the reliability of this questionnaire, its value as a measure of non-academic accomplishment would be very dubious indeed. For this reason, a random sample of twenty-four pupils who were available eleven months after the initial testing programme, was given the NAAQ again. The rank order correlation between the initial and subsequent total score was .807 ($p < .0005$, one-tailed). Considering the wording of the questionnaire which is not strictly objective and open to many interpretations, as well as the interval of eleven months between the two administrations, the obtained reliability may be considered high enough to continue the use of this questionnaire as a source of information about the pupils' accomplishments outside the academic field. Also, with further try-out and psychometric refinement, it may reach a higher reliability level than obtained in the present study.

In view of the diverse nature of the activities covered by this questionnaire it is not surprising that Table 2.2 in Appendix A shows low, and often non-significant inter-correlations.

It will however be noticed that with the exception of Music, Science and Sports (columns 5, 6 and 9) all other activities have a small but highly significant correlation with the NAAQ total score. As the note in that table points out, these correlations are between completely independent scores and yet the fact that five of them are highly significant ($p < .001$) one is significant at .019, one at .067 and one at .095, shows that with the exception of Sports, the questionnaire as a whole has a certain coherence in the sense that those individuals who indicate an involvement in one kind of activity are also involved in some other non-academic pursuits.

A look at some of the correlations in Table 2. 2 points to a general trend towards higher correlations between areas which seem related in terms of the type of activity they cover. For example, the highest correlation ($r = .317$ $p < .001$) is between accomplishments in Writing and Debate, both presumably requiring an interest and facility with verbal material. Similarly, the next highest correlation is between Drama and Debate ($r = .272$, $p < .002$) both being fields of activity requiring expository skills. The next pair yielding a correlation of .251 ($p < .004$) is that of Social Service and Leadership, again both having much in common since they involve dealing with people. Yet another significant relationship in Table 2.3 is that between Drama and Leadership ($r = .219$, $p < .012$), areas requiring the ability to adopt and play certain roles convincingly enough to influence others. Thus, it does seem that despite its rough and ready form, the NAAQ has a reasonable amount of consistency, at least as an exploratory research instrument.

Like the scores on the Divergent Thinking Test, raw score statistics of the NAAQ also showed considerable skewness (1.457) towards the lower end of the scale. The distribution was therefore normalised to a mean of 10 and a standard deviation of 5 so that parametric statistics may be used as with other measures.*

---

*Rock, Evans and Klein (1969) have suggested that when skewed distributions of this kind are obtained it may still be possible to use parametric statistics on sub-groups formed on the basis of some moderator variable which sorts the individuals in such a way that the within-group distribution is close to normal. However, the design of the present study ruled out such a procedure.

SECTION THREE

STATEMENT AND DISCUSSION OF RESULTS

CHAPTER VII

ANALYSIS OF DATA FOR THE WHOLE SAMPLE

Table 1.1 in Appendix A gives means and standard deviations for each school class as well as for the whole sample (last column). It also gives significance of differences between pairs of class means. Other tables in Appendix A (Tables 1.2 to 1.6) give the same statistics for Boys/Girls; Leavers/Non-Leavers; High, Middle and Low levels of IQ; Top, Middle and Leaver's sets formed by combining school classes. The next series of tables (Tables 2.1 to 2.5) are tables of correlations. The purpose of these tables is to present in one place, the basic statistical information regarding the different variables under consideration in this study. Reference will be made to certain parts of these tables at the appropriate places in the discussion of results.

Distribution of Raw Scores

Since parametric statistics are being used for most of the analyses, it is necessary to consider the shape of the distribution of different variables for the whole sample. As mentioned before (p.162) IQ was normally distributed the coefficient of skewness being 0.229, which was not significantly different from zero. Teachers' marks for English and Arithmetic were taken from school records and rescaled to the IQ mean and standard deviation using the NFER method described by Yates and Pidgeon (1957). Consequently, the distributions of these two variables were also almost the same as that of IQ. Of the remaining six variables (DT, MI, nAch, NAAQ, SEB and Career Choice) the two which showed considerable deviation from normality were DT (skewness 1.363) and NAAQ (skewness 1.457).

DT raw scores were normalised to a mean of 110 and a standard deviation of 11 to make them comparable to IQ and the two attainment measures (English and Arithmetic). Raw scores for NAAQ were also normalised to a mean of 10 and a standard deviation of 5. These figures were chosen for the mean and sd of NAAQ as they were more representative of the actual level of non-academic achievement reported (raw score m = 7.130, sd = 6.400) than the IQ mean and sd were to which DT scores had been normalised.

Correlation Between IQ and Divergent Thinking

Before we look in detail at IQ-DT correlations for different levels of IQ, it is necessary to find out how these two variables are correlated for the whole range of scores. Table 2.1 (Appendix A) gives the value of IQ-DT correlation for the whole sample as .467 ($p<.001$). This degree of relationship between IQ and divergent thinking is not too different from that reported in other British studies (Bennett 1973, Lytton and Cotton 1969) although it is relatively higher than the correlations found in American ones. The chart in Appendix D summarises the main findings of these studies. Information regarding IQ-DT correlations is given in column 14 of this chart. It shows that Haddon and Lytton (1968) obtained a correlation of .480 between VRQ and DT for a sample that covered a fairly normal range of VRQ. Haddon and Lytton (1968) do not report the actual standard deviations for their samples but the mean VRQ of their samples was also close to the population average (Formal schools m = 101.75, Informal schools m = 101.14). However, the IQ-DT correlations in the Haddon and Lytton study and the present one are not strictly comparable as the mean IQ in this study is 110.296 with a standard deviation of 11.264. According to the threshold hypothesis the present study should have produced a lower IQ-DT correlation as the mean IQ is almost one standard deviation above the population average of 100.

In fact Hasan and Butcher (1966) did find a much higher correlation (r = .743) between their creativity aggregate and VRQ with a sample that had a mean VRQ of 102 and a standard deviation of 12. And, in a follow-up study of the pupils they had tested at primary school, Haddon and Lytton (1971) also found a higher correlation (r = .615) between VRQ at 11 and DT scores at 15. This increase in correlation is remarkable for two reasons: Firstly, there was a time lag of nearly four years between obtaining the two measures correlated; and other studies of the relationship between cognitive abilities over time have reported lower correlations between ability and attainment measures as the time lag increases, (France 1964, Nisbet and Entwistle 1969). Secondly, in the follow-up study Haddon and Lytton could trace only 151 of the 211 subjects tested earlier. Again, standard deviations for VRQ or any other variable are not reported, but it is possible that loss of subjects may have narrowed down the spread of VRQ. In such a situation lower correlations would be expected in the follow-up study.

In Kogan and Pankove's longitudinal study referred to earlier (p.16) there was a similar trend for IQ-DT correlations to increase at the secondary school level. It may be that as pupils move from primary to secondary schools they become less free in associative thinking of the type required for doing well on open-ended tests. They may also become more cautious, self-critical and rational in evaluating their own responses to such tests. In other words there may be more "self-censoring" going on, as Heim (1970) has suggested. Hence the higher correlations in follow-up studies between VRQ and DT. If this is the case then the high IQ-DT correlation in Hasan and Butcher's (1966) study can also be explained in terms of the age of their subjects who were in the second year of secondary education in Scotland. Thus, they were 13 years old, whereas the subjects in Haddon and Lytton's (1968) earlier study were in their last year at primary school in England.

Admittedly the actual difference in age in the two studies* is not great, but when this is combined with the difference between a primary school ethos and a secondary school ethos, the effect on how free children feel to express themselves in an unconventional test is perhaps more significant. For example, Nisbet and Entwistle (1969) have argued that "the traditional transfer procedure does provide an unnecessarily sharp break for most children". In evidence they quote Murdoch's (1966) analysis of essays written by children six weeks after transferring to secondary school, describing how they felt about the transfer. According to this analysis "only about 10 per cent of the children in this sample appeared to find transfer wholly enjoyable". Commenting on the implications of this finding Nisbet and Entwistlestate: "The composite description is full of disturbing and frightening experiences... The feelings of insignificance and bewilderment wear off, but our other research shows that the after-effects seem to leave their mark on

* Hasan and Butcher (1966), Haddon and Lytton (1968).

children's academic performance throughout the first year at secondary school. First-year performance influences a child's attitudes which in turn influence second year performance" (p.87). If this is the effect of transfer on conventional academic attainment, the negative effects on divergent thinking may well be greater.

Another piece of evidence regarding the effect of age on divergent thinking (and its correlation with IQ) comes from Torrance (1965). In a number of cross-cultural studies of divergent thinking at different age/grade levels, he found a sharp decline in DT scores around the age of 13. Such a developmental trend may also account for the higher IQ-DT correlation generally reported at the secondary school level. Thus, a correlation of .467 obtained in the present study is in the expected direction when we consider that it is for a sample that is considerably superior in ability (m = 110, sd = 11). With a more representative sample, as in Hasan and Butcher's study, it may well have been in the sixties or seventies. This degree of over-lap between IQ and DT raises further doubts regarding the existence of a distinct convergent and divergent cognitive style as postulated by Guilford, Torrance, Getzels and Jackson and Wallach and Kogan.

Nor can the present IQ-DT correlation be explained away in terms of the lack of convergent and discriminant validity of the divergent thinking tests as Wallach (1970) has argued was the reason for the high IQ-DT correlation in Hasan and Butcher's study. It has been mentioned before (p.162) that the trend of item-total correlations of DT and the average correlations between DT items and IQ, English, Arithmetic shows that in the present study this is not the case. The average correlations computed from the matrix in Table 2.3 (Appendix A) are as follows:

Table VII.1 Average correlations of DT items with Flexibility Total, Originality Total, IQ, English and Arithmetic

	Flexibility Total	Originality Total	IQ	English	Arithmetic
Average r of DT items	.60	.52	.30	.37	.27

It will be recalled that these correlations are between independent scores; that is to say, the item-total correlations are not artificially inflated because of a part-whole relationship. Scores on the individual items were removed from the total before these correlations were computed. Thus, all correlations in the above table are comparable. This procedure for determining the convergent and discriminant validity of DT tests also meets Guilford's (1968) objection to the generally used method of comparing correlation between individual divergent tests with a composite IQ. According to Guilford, the latter procedure is invalid because being a composite measure, IQ is able to share more sources of variance with individual divergent tests than they can share between them. The correlations given in Table VII.1 above overcome this difficulty as they are all between individual items and a total score. As a further check on whether the different DT sub-tests do hold together internally, average inter-correlations among sub-test items were also computed. These were

.51, .57 and .53 for the Uses, Similarities and Consequences subtests respectively. Thus, in view of a reasonable degree of convergent and discriminant validity which the divergent thinking tests possess in this study, and the relatively superior ability of the sample, a correlation of .467 between IQ and DT implies a much closer relationship between the convergent and divergent modes of thinking than would have been expected on theoretical grounds.

Inter-Correlations Among IQ, English, Arithmetic, Career Choice and SEB

Correlation Between IQ-English and IQ-Arithmetic:

Table 2.1 (Appendix A) shows that with one exception (English-Arithmetic correlation to be discussed later), correlations between IQ-English and IQ-Arithmetic are the highest (r =.679 and .632 respectively). As these correlations are based on the rescaled English and Arithmetic rank-orders it may be that they have been artificially inflated by the particular rescaling procedure used, and are therefore not strictly comparable to other correlations between truly independent scores.

On the basis of the mathematical reasoning underlying the rescaling procedure used here (which involves rescaling on the major axis of the correlational ellipse) Pilliner (1958, 1965) has shown that the correlation obtained in this was "is in part spurious". This is because the method used (Pilliner's Method 3) equates the class means of the teachers' estimates after rescaling, to the class means of the rescaled test, thus introducing a correlation of unity between the two sets of class <u>means</u>. However, provided that the class means on the rescaling test are not unduly wide-spread, the overall correlation between rescaled assessments and rescaling test is raised only slightly above the average within-class correlation.

That the present correlations are not much higher than could have been obtained between truly independent variables is further suggested by the high value of VRQ-English (r = .845) and VRQ-Arithmetic (r = .734) correlations found by Hasan (1965) when objective attainment tests were used instead of rescaled marks. Nisbet and Entwistle (1969) also report finding an almost identical pattern of correlations between a verbal reasoning test and a criterion rescaled on it, and that between the verbal reasoning test and on unscaled criterion. This may be taken as further evidence that the rescaling procedure used in the present study is unlikely to have artificially raised the correlations between rescaled scores and the test used for rescaling.

The question of the validity of a single set of examination marks as an index of English or Arithmetic attainment as such, was also considered in some detail before these marks were used. From school records English and Arithmetic marks from four previous school examinations (December and June examinations during the first and second year of secondary education) were averaged and ranked, for a random sample of pupils. These rank-orders were correlated with the third year December examination rank-orders to estimate the extent to which the most recent set of marks corresponded with these pupils' previous attainment. The obtained correlations were .900 for Arithmetic and .899 for English both significant beyond the .001 level.

Correlation Between English and Arithmetic:

A product-moment correlation of .704 between English and Arithmetic (Table 2.1, Appendix A) suggests that not only is performance in these subjects fairly consistent over time, but also that on the whole, those who do well on one of these tend to do well on the other too. Again, it is possible that the rescaling procedure has slightly raised the correlations but the difference made is not likely to be much as Hasan (1965) has also reported a correlation of .659 between attainment tests of English and Arithmetic. In that study no rescaling on a common test was carried out because objective attainment tests had been administered as part of the research.

The validity of using performance in English and Arithmetic alone as an index of academic achievement in general was also considered before deciding to confine the study to these two measures of attainment. It was reasoned that since tests of verbal and numerical ability make up the greater part of most intelligence tests, and measured intelligence has been shown to be positively related to academic performance at least at the secondary level (Vernon 1957), attainment in English and Arithmetic is also likely to be a fairly representative measure of the general level of academic performance. More recently, Nisbet and Entwistle (1969) have also provided some empirical evidence for such a line of argument. In an extensive follow-up study of school-children from the primary to secondary stage, they found that objective tests of English and Arithmetic, as well as teachers' scaled estimates of these subjects taken at age 12 had correlations ranging from .787 to .841 with a scaled criterion of "secondary school attainment during the second year" (age 14). It seems therefore, that English and Arithmetic attainment, whether obtained from standardised tests or teachers' estimates, are a reasonably valid index of the general level of academic achievement in secondary school.

Correlations Between IQ, English, Arithmetic and Career Choice/SEB:

Another set of correlations in Table 2.1 (Appendix A) which show a highly significant positive relationship are those between Career Choice and IQ ($r = .543$), English ($r = .526$), Arithmetic ($r = .494$). The obvious interpretation of these correlations is that, for the present sample, ability, attainment and aspiration are quite closely related. Not only does academic achievement at school share nearly half of its variance with IQ but IQ and English/Arithmetic attainment also seem to be the best (not the only) measures on which to base predictions about these pupils' level of achievement (in terms of the type of careers they will have) after completing formal education.

Of course, it is quite likely that some individuals in the present study may not actually end up doing what they gave as a career choice at school, but it is also unlikely that on the whole the change will be so great as to alter the value of these correlations to any considerable degree. For example, a boy who said he wanted to be a painter and decorator is not likely to change to such an extent as to end up being a research chemist or vice-versa. Certainly people will move up and down the occupational ladder a little, but even after such a movement there is every chance that the correlation between IQ and career choice will remain either constant (due to the compensatory effect of change in both directions - up and down) or it may actually rise (due to a better "fit" between ability and aspiration as time goes on).

A study of the Scottish Council for Research in Education (SCRE) published in 1970 throws some light on the question of how far school-leavers succeed or fail to realise the occupational goals they set for themselves while at school. This study was designed to find out what changes occur in the basic skills of English and Arithmetic, job preference, leisure-time interest and activities etc., between the last year of secondary education and a year after this time when the young people have been working for a year. Commenting on the extent of divergence between job preference given when at school and actual occupation, the report states: "77 per cent of the group either realised their occupational ambitions or were very close to it (being one category above or below)" p.106. The SCRE study also looked at the question of how stable the job preference expressed at 15 is. Not counting the pupils who were still at school at the time of the follow-up study, the investigation found that 62 per cent gave the same preference after a ten-month period. Quite plausibly they point out that had the pupils still at school been included in this analysis, "the percentage would doubtless have been higher".

Another part of the SCRE report relevant for the present discussion is the "almost perfect correspondence between the hierarchicalstructure of "job preferred" and achievement as measured by the tests... It would appear that pupils at the age of 15 are well aware of their proficiencies and deficiencies when looking forward to the kind of job that will suit them when they are grown-up" (p.94-95). No correlations or other measures of association between attainment and job preference are given in the SCRE report and so direct comparisons with the findings of the present study are not possible. However, the data of this much larger investigation do provide considerable evidence for the empirical validity of the correlations found between Career Choice and English and Career Choice and Arithmetic in the present study.

Turning to the correlations between IQ, English, Arithmetic and SEB, we see from Table 2.1 (Appendix A) that although all three are positive and significant the values are in every case lower than those for Career Choice discussed above. It will be noted that in the research design of the present study, SEB is to be considered as an independent variable, along with IQ, DT, MI and nAch, in the regression analysis to be reported later on. In view of the small proportion of common variance between SEB and IQ/DT and the near zero correlations of SEB with MI and nAch, this appears to have been a reasonable decision. For, in a post hoc manner, it can be seen that, with the exception of the IQ-DT correlation, there is very little over-lap among the other independent variables. Under these circumstances, whatever contribution they made to the regression equation could be considered to be relatively unique and free from the confounding effects of a third variable. Conceptually too, there was reason to expect that SEB would make a significant contribution to non-academic accomplishments (NAAQ) by providing greater opportunity for participation in such activities and also by encouraging their pursuit in positive ways. In an earlier study, Skager , Shultz and Klein (1965) have shown that the Quality score of their non-academic accomplishments questionnaire was significantly related to a measure of intellectual stimulation in the home.

Correlations of DT with English, Arithmetic, NAAQ, SEB and Career Choice

Having considered the pattern of relationships between IQ and the other variables (with the exception of MI and nAch to be discussed later) we can now turn to the correlations between DT and the same variables to see if there is any marked difference in the pattern of correlations here. We have already seen that IQ and DT are themselves positively correlated ($r = .467$ $p < .001$), so the difference in relationships is not likely to be too much. In the discussion below DT-English and DT-Arithmetic correlations will be considered first, followed by DT-NAAQ and DT-SEB.

Correlations Between DT-English, DT-Arithmetic:

Table 2.1 (Appendix A) shows that like IQ, DT also has highly significant positive correlations with English ($r = .539$) and Arithmetic ($r = .383$) but the difference between DT-English and DT-Arithmetic correlations is much more than was the case with IQ. This is according to expectation since the divergent thinking tests used in the present study were all verbal, whereas the intelligence test had some numerical items also and is on the whole a test of reasoning and logical thinking ability. It is therefore not surprising that it correlates about equally with English and Arithmetic attainment. Also, as an intelligence test it is a test of general ability which accounts for the overall level of achievement that people actually show and it is again reasonable to expect that such a test will show a similar relationship to the main areas of academic achievement. In comparison with this, the divergent thinking test instructions explicitly stated that it was a test of imagination, and, as mentioned before, the "items" were all verbal. What is surprising in these correlations is the fact that divergent thinking tests of a purely verbal type should correlate with Arithmetic attainment. But then we have seen that IQ and DT are themselves related, so this may to some extent account for the correlation between DT and Arithmetic.

There is some evidence from earlier studies too that verbal tests of divergent thinking do tend to have higher correlations with measures of language attainment than they do with arithmetic. Cicirelli (1965) found that factors of verbal fluency, originality and elaboration had higher correlations with reading than with arithmetic, while the non-verbal factors had about the same correlations with these criteria of attainment. Similarly, Clark, Veldman and Thorpe (1965) have reported that their "high divergent" group had significantly higher scores on Reading and Word Fluency Tests. In a study by Yamamoto and Chimbidis (1966) this trend was again confirmed. Hasan (1965) too reported a similar finding in her earlier study, not only for total score correlations between Creativity Aggregate and English/Arithmetic Quotients but also in the correlations between individual creativity tests and these quotients. More recently, Hargreaves and Bolton (1972) also found that DT had a correlation of .42 with Mednick's (1962) Remote Associates Test, another verbal test. Good performance on this test would appear to depend much on vocabulary in particular. So it seems as if the contribution of divergent thinking to academic achievement is likely to be more subject-specific than that of IQ, since most divergent thinking tests are of a verbal type.

Correlations Between DT and NAAQ:

Along with the IQ-DT correlation, that between DT and Non-academic Accomplishments Questionnaire is perhaps of most interest in the present study. Table 2.1 (Appendix A) shows that of all the correlations between NAAQ and other variables, DT-NAAQ correlation is the highest ($r = .371$ $p < .001$). This provides some support to Wallach and Wing's (1969) argument regarding the use of DT scores for predicting non-academic accomplishments. In view of Werts' (1967) criticism regarding the unsuitability of correlational analysis in the studies of Holland et al, Wallach and Wing do not themselves give any correlations between their measures of divergent thinking and non-academic accomplishments; they use extreme group, mean score comparisons instead. But, for the predictive validity of divergent thinking for NAAQ to be clearly established, it should also be shown that there is a significant positive relationship between DT and non-academic accomplishments over the whole range of ability. For this purpose, the relevant correlations in the present study are in the expected direction: the correlation between IQ and NAAQ is .210 ($p < .05$), that between DT and NAAQ is .371 ($p < .001$). It will be recalled that, unlike the sample in Wallach and Wing's study, the subjects in the present one were not a volunteer group, and although above average in ability, they were still not as highly selected as the University freshmen in the American study. The test-retest reliability of .807 ($p < .0005$) for NAAQ in the present study, also mentioned before, suggests that the NAAQ scores have reasonable stability. Considering all these factors together, it does seem that inspite of considerable overlap between them, IQ and DT are differentially related to NAAQ.

Correlations Between DT and SEB/Career Choice:

In the present study, the correlation between DT and SEB was .211 ($p < .05$) and between DT and Career Choice .414 ($p < .001$). The difference between these correlations is in the same direction as it was in the case of IQ, which may follow partly from the positive correlation between IQ and DT itself. Not many studies have looked at the relationship between divergent thinking and socio-economic status or background directly; although Getzels and Jackson (1962) did find that fathers of the "High Creatives" were mainly from the business class and those of the "High IQ" came from such professions as University teaching, research or editing. More recently, two studies in Britain did consider this question, but the findings were in opposite directions. In the earlier study, Lytton and Cotton (1969) found a correlation of .26 between DT and social class, determined by father's occupation. In the same study, the correlation between VRQ and social class was .58. On the basis of these correlations the authors suggested that divergent thinking tests were "less dependent on environmental influences". In view of the following considerations this conclusion seems hardly warranted:

(a) A number of studies have shown that DT scores are quite susceptible to fluctuation with changes in testing conditions/ test instructions (Boersma and O'Bryan 1968, Datta 1963, Dentler and Mackler 1964, Kogan and Pankove 1972, Parnes and Meadows 1959, Ward, Kogan and Pankove 1972).

(b) Other studies have also shown that DT scores vary according to the permissiveness and authoritarianism of the subjects' environment (Haddon and Lytton 1968, Marino 1971, Ogletree 1970, Walker 1967), although the findings in this respect are rather mixed. Haddon and Lytton and Ogletree report finding higher

DT scores in permissive, informal schools; whereas Walker found that pupils in such schools did not do as well on DT as pupils from more formal schools. In "Cross-national Comparisons of Catholic-Protestant Creativity Differences", Marino found that according to prediction, Protestant students in U.S.A. and N. Ireland had significantly higher DT scores, but there were no differences between the groups in Eire and and Scotland.

(c) Getzels and Jackson (1962) have also pointed out in some detail the consistent differences found between the family environment of their "Highly Creative" and "Highly Intelligent" group, which suggests that these environmental differences may be related to a particular mode of cognitive functioning.

In their follow-up study Haddon and Lytton (1971) also looked at the correlation between Socio-Economic Status (SES) and DT/VRQ. They did not find the type of difference between SES-DT and SET-VRQ reported by Lytton and Cotton earlier. In fact, the trend of their correlations was very similar to that of the present study: DT and VRQ correlated with SES .351 and .395 respectively, thus showing only a very slightly higher correlation between VRQ and SES than between DT and SES. Haddon and Lytton explain this difference between the two findings in terms of the more selected nature of Lytton and Cotton's sample, which had a mean VRQ of 111 (Haddon and Lytton's had 101.9) and a sizeable proportion of pupils were from a direct grant grammar school, which according to Haddon and Lytton (1971) "implies a disproportionately high number of high status parents" (p.143). It is not clear how these differences account for the differential correlations in

Lytton and Cotton's (1969) study; for if it is narrowing in the range of SES that is implied, then it ought to affect both (VRQ-SES and DT-SES) correlations and if it is selection in VRQ that is crucial then this ought to produce a lower, or about the same, value of r for VRQ and DT, rather than the direction of difference reported by Lytton and Cotton.

Correlations Between Motivation Index (MI) and Other Variables

The generally low correlations between MI and the other variables (Table 2.1, Appendix A) suggested that there may be a curvilinear relationship involved, and therefore that the product-moment correlation being computed was not a suitable measure of association between MI and these variables. A test of linearity, using the analysis of variance procedure (McNemar 1962), was applied to check on this possibility. For this purpose MI was considered as the criterion variable and all other variables were divided up into four levels at roughly one standard deviation below the mean, one standard deviation above the mean and the scores beyond this range at both ends of the scale. The mean, standard deviations of MI and F ratios for the test of linearity are presented in the table below:

Table VII.2 : Means, Standard Deviations, F of MI at Four Levels of IQ, DT, nAch, SEB, English, Arithmetic, NAAQ and Career Choice.

Criterion Variable		Independent Variables				F
		IQ Score Range				
		80-100	101-110	111-120	121-140	0.745
MOTIVATION INDEX	n	18	36	37	17	ns
	m	101.611	114.472	111.081	112.824	df
	sd	25.433	33.166	31.409	34.390	2,104
		DT Score Range				
		80-100	101-110	111-120	121-140	8.690
	n	18	41	29	20	**
	m	100.333	120.561	95.517	122.950	df
	sd	22.476	24.644	38.238	29.547	2,104
		nAch Score Range				
		0	1-3	4-6	7-9	1.267
	n	44	14	25	25	ns
	m	97.159	119.000	118.760	122.720	df
	sd	26.806	26.218	32.169	33.588	2,104
		SEB Score Range				
		1-2	3	4-5	6	4.965
	n	10	31	36	13	**
	m	96.600	110.065	123.750	97.077	df
	sd	23.899	33.822	31.255	31.434	2,86

*p<.05 **p<.01

cont.

Table VII.2 : Cont.

Criterion Variable		Independent Variables				F
		English Score Range				
		80 - 100	101-110	111-120	121-140	
MOTIVATION INDEX	n	18	37	35	18	1.380 ns
	m	97.444	108.676	119.629	112.000	df
	sd	27.866	28.883	33.942	32.062	2,104
		Arithmetic Score Range				
		80-100	101-110	111-120	121-140	
	n	18	36	34	15	3.992 *
	m	104.833	118.778	100.853	118.933	df
	sd	31.444	25.802	32.825	32.110	2,99
		NAAQ Score Range				
		0-5	6-10	11-15	16-21	
	n	28	39	28	13	1.613 ns
	m	118.107	110.744	101.571	116.000	df
	sd	30.454	31.831	34.873	21.350	2,104
		Career Choice Score Range				
		0-3	4	5	6	
	n	15	21	28	13	1.048 ns
	m	103.800	114.810	117.571	106.538	df
	sd	35.762	36.235	26.160	34.705	2,73

From the above table it can be seen that although there is a slight drop in MI mean scores at the higher levels of IQ, the test of linearity showed an F ratio of 0.745. The required F for rejecting the hypothesis of linearity of regression is 4.82 ($p < .01$). It therefore seemed reasonable to conclude that the deviation from linearity shown in mean scores was not statistically significant, and curvilinearity would not account for the low correlations obtained between IQ and MI.

It will be recalled that Arnold's (1962) scoring scheme was used for obtaining the Motivation Index from TAT protocols. Arnold herself has reported correlation ranging from .468 to .582 between MI and various measures of general intelligence. The present finding of an almost zero correlation between IQ and MI ($r = .078$) is therefore contrary to expectation. There seem to be two possible explanations for it: Firstly, the low intra-scorer and inter-scorer reliability found in the present study (rank-order correlations .628 $p < .005$ and .586 $p < .005$ respectively) may to some extent account for the low correlations. Arnold gives inter-scorer reliability in terms of percentage of agreement between two scorers, rather than as a correlation between scores. It is therefore not possible to make any direct comparison in this respect. The percentage of agreement in studies reported by Arnold ranges from 80 to 97. But, it must be noted that these percentages are for an earlier version of the scoring scheme, according to which protocols were simply scored plus or minus, rather than on a five point scale as is the case in the final scoring scheme used for this study. The earlier system inevitably raises the probability of agreement by chance alone, since there are only two categories to which a story can be assigned, instead of the five categories in the later system. Another factor that may have raised the percentage agreement in these earlier studies is that they were conducted not with representative samples but with groups pre-selected on such variables as effectiveness as teacher/non-effectiveness as teacher or for being offenders/non-offenders etc. This excluded subjects in the middle range of these variables. Again, the chances of agreement are greater when extreme group comparisons are used and this may also be responsible for the high percentage of agreement between judges reported by Arnold.

The second possible explanation for the near zero correlation between IQ and MI is also related to the question of reliability in terms of the number of stories on which MI is based. Arnold recommends a minimum of ten stories per person, but in the present study, only six slides were used, due largely to the limitation of testing time available. Aslo, earlier research (Lindzey and Heinemann 1955, Reitman and Atkinson 1958) has suggested that increasing the length of protocols beyond four cards or increasing the response time beyond five minutes per story does not improve reliability. As far as the present study is concerned, in retrospect it appears that asking for six stories was about the right number to engage and sustain the interest of pupils. Judging from the brevity and stereotyped quality of most protocols, it also seems as if the task was a rather strange and uninspiring one for most of them. However, theoretically it is possible that increasing the number of stories would have improved the reliability of the MI, which in turn may have produced higher correlations, not only with IQ but with other variables too.

Of course, the simplest and most straightforward interpretation of the low correlations between IQ and MI is that, in this sample there was no significant relationship between IQ and achievement motivation measured by the story sequence analysis procedure. On the other hand, it is also true that although a high correlation may be taken as evidence of a significant relationship between two variables, a low correlation does not necessarily prove the absence of such a relationship. On this argument it may well be that the slight curvilinearity and low reliability discussed above are confounding the actual relationship between the two variables involved here. Therefore, inspite of this low correlation with IQ, MI was retained in the rest of the analysis because it was still possible that, as an independent variable it may make a significant contribution to actual academic or non-academic achievement.

It can be seen from Table 2.1 (Appendix A) that MI has a correlation of .331 ($p < .001$) with nAch, which suggests that it has some validity as a measure of achievement motivation. Admittedly, the predictive validity of nAch itself for actual achievement in life has not been clearly established, but some of the early studies by McClelland and his colleagues did report a significant relationship between nAch and real-life achievement (Atkinson 1958, Crocket 1962, Douvan 1956, McClelland et al 1953, Rosen 1956, Veroff 1961).* As the present study was concerned not only with academic but

*Referring to some of these studies, Murstein (1963) has rather cynically remarked that nAch scores have been used to predict "even the rise and fall of empires.".

also with non-academic achievement it was expected that if MI has any validity, it would show up in the form of a significant contribution towards the latter kind of achievement. Since Arnold's (1962) scoring scheme is based mainly on the distinction between a positive and negative attitude to life rather than on the notion of "competition with a standard of excellence", as nAch scoring is; the expectation was that MI would be positively related to achievements initiated by pupils themselves, whereas nAch would be more closely associated with academic achievement, a field in which standards of excellence are clearly defined and there is more explicit competition.

From Table 2.1 (Appendix A) it can be seen that, with the exception of nAch, MI had near zero and non-significant correlations with all other variables. Results for the test of linearity presented in Table VII.2 above also show that there are only two variables (DT and SES) for which the hypothesis of linearity of regression can be rejected with any confidence. Therefore, curvilinearity would not entirely account for those low correlations. Rather, it seems as if in the present study, MI has no predictive value either for academic or for non-academic achievement.

Correlations Between nAch and Other Variables

In comparison with the mainly negative finding regarding the Motivation Index discussed above, Table 2.1 (Appendix A) shows that nAch has a significant positive correlation with IQ, DT, English and Arithmetic, although the value of correlations is not very high, ranging from .205 (IQ-nAch) to .274 (English-nAch), all significant at the 5 percent level or better. These correlations are in the expected direction, for if nAch as any validity in terms of actual achievement, it is likely to be positively related to IQ as well as academic attainment, even if the correlations are only modest in value. There is no previous research with which to compare the DT-nAch correlations found in the present study, but considering that IQ and DT are positively correlated themselves, the similaritiy in the value of nAch correlations with these variables is also accrording to expectation.

With reference to IQ-nAch correlations in the existing literature, Robinson has noted: "Little interest has been shown in possible relationships between intelligence and achievement motivation by American workers". The few studies which did look at the question, reported rather mixed findings (French 1958; Lowell 1953; McClelland 1953, 1958). It seems a serious omission to have overlooked this very fundamental question, for it is possible that some of the differences in actual achievement found in validity studies of nAch could also have been explained in terms of IQ differences. For example, in his celebrated study of the socio-cultural origins of achievement motivation, Rosen (1956) found that a significantly greater percentage of high nAch scorers got average school grades of "B" or above; and similarly, more of the low nAch scorers got average grades of "C" or below. From this, and a number of other analyses, using social class and value orientation as other independent variables, Rosen concluded: "this study reveals a significant relationship between achievement motivation and grades, and between values and educational aspirations". He admits that "Achievement motivation and achievement oriented values are not, of course, the only factors related to academic success and educational aspiration" (p.506-507). But, only in passing does he mention the possible effects of intelligence on the differences he found. As Lavin (1965) has pointed out a number of other studies which have argued for the predictive validity of nAch for academic achievement, must also be considered inconclusive because they failed to control for intelligence.

However, recently some British workers have looked at the relationship between nAch and measured intelligence, and the findings have shown positive correlations ranging from .2 to .4. It should be noted that some of the American studies referred to earlier had reported a similar trend*. In a longitudinal study, Kagan and Moss (1959) also found a positive relationship between the nAch scores of 8½ year olds and the increase in their IQ between the ages of 6 to 10. In fact, as early as 1940, Harrison had shown that estimates of IQ derived from TAT protocols had a correlation of .78 with actual IQ. More recently, Henry (1956) has also reported a slightly higher correlation (r = .85) between estimated intelligence from TAT stories and measured IQ.

In view of these earlier studies, it is not surprising that Bruckman (1966) and Jayasuria (1960) have argued that IQ is the most crucial variable in determining nAch differences, and when its effects are removed either statistically (Bruckman) or by sample selection (Jayasuria), social class and school stream differences

---

*In Lowell's study the correlation between nAch and ACE was .28 (n=37) and McClelland found a correlation of .42 (n=30) between nAch and SAT scores. (McClelland et al 1953.) No significance levels are given but in view of the small numbers the correlations may just reach significance at the 5% level.

related to nAch become insiginificant. Implicit in this relationship between IQ and nAch is the possibility of a positive feed-back effect on nAch from "success experiences", which in turn may themselves be related to IQ. For example, high IQ pupils are also likely to be doing well at school. This experience of success may reinforce their desire to do well, thus leading to a high need for achievement which is reflected in their nAch scores. Also, even within social classes, the more intelligent child will probably have received more positive reinforcement from his mother and other members of the family during early childhood. According to McClelland's theory of the origins of achievement motivation in childhood experiences of this kind, a child who had received greater rewards in achievement situations is likely to have a higher level of motivation. Thus intelligence would be related to nAch through success experiences.

Bruckman (1966) and Jayasuria (1960) had offered this line of reasoning as a possible explanation for the association they had found between IQ and nAch. In their studies, Robinson (1964) and Finlayson (1972) specifically included certain measures of success in school as moderating variables and found support for this view. However, Table 2.1 (Appendix A) shows that in the present study English and Arithmetic had low correlations with nAch (.274 and .245 respectively) which suggests that the association between success (judged by attainment scores in this case) and nAch is not very strong in this sample.

The correlations of nAch with IQ, DT, English and Arithmetic discussed so far are in the expected direction. But, when we look at the other correlations (nAch with SEB, NAAQ and Career Choice) none of them is significantly different from zero. The lack of a significant relationship with SEB requires some explanation as earlier studies in this field have reported a positive relationship between middle-class, child-rearing practices and the development of achievement motivation (Douvan 1958, Winterbottom 1958). But, as Bruckman (1967) and Jayasuria (1960) have shown when ability is controlled for, the association between social-class and nAch no longer remains significant. In fact, with quite a large sample which included pupils from A, B and C streams in Secondary schools, Bruckman also found only low correlations between SEB and nAch (for boys $r = .14$ $p < .05$ $n=204$; for girls $r = .20$ $p<.01$ $n=179$). Therefore it is not surprising that in the present study the SEB-nAch correlation is nearly zero when we consider that the correlation of nAch with IQ is itself very modest ($r =.205$, $p<.05$). Also, it seems now that the expectation of any consistent relationship with SEB, when SEB is measured on a 6 or 7 point scale (as is the case in most studies), is taking a rather over-simplified view of the situation. Swift (1966, 1968) has suggested that it is better to conceive of social classes as "social environments" in which there is a complex interaction of values, attitudes and aspirations which may cut across classes. In these circumstances the linear association between school attainment and high achievement motivation or social class will not be obtained if social classes are treated as discrete entities.

The near zero correlation of nAch with NAAQ is more difficult to explain. It was expected that if the concept of achievement motivation, measured by the analysis of imaginative stories has any validity for predicting real life achievement, it would show at least a moderately positive association with such a general index of non-academic accomplishments, as the NAAQ. According to McClelland's theory, the achievement motive is an accumulation of specific habits learned in early childhood along the principles of reinforcement. For this reason, de Charms (1968) has argued that by analysing TAT stories for achievement motivation we "sample the extensity of the thoughts about achievement... Measurement based on the concept of extensity aims at generality rather than specificity and results should relate to general dispositions and to behavioural trends over time rather than to specific behavioural patterns" (p.227).

It will be recalled that the NAAQ covers a fairly varied range of activities over a three year time span of secondary education. Also, it is concerned with activities which pupils are likely to have undertaken as a matter of choice, thus getting involved in a situations of "competition with a standard of excellence". It would appear then that nAch as a measure of motivation ought to relate positively with NAAQ as a measure of real-life achievement. We have also seen that NAAQ score may be considered a fairly stable index of such achievement (test-retest reliability over a eleven-month period was .807). The fact that the expected positive relationship between nAch and NAAQ was not obtained suggests two possibilities: Firstly, the moderate, positive correlations of nAch and IQ, DT, English, Arithmetic may be interpreted not so much as a measure of how closely need to achieve is related to actual attainment, but rather, as an indication of the extent to which the ability to write a story around a fairly ambiguous picture may be involved. To use the distinction traditionally made in psychology, a congitive rather than an affective or motivational variable may be the basis of those positive relationships. Hence, the lack of any significant association between nAch and non-academic achievement as measured by the NAAQ. Secondly, fantasy measures of achievement motivation such as the nAch score may just not be related to what people actually do.

It can be seen that the two possible explanations offered above are in fact related. If fantasy measures of achievement motivation are not related to real-life achievement then the low positive correlations obtained in this study between nAch and academic attainment may be accounted for in terms of ability rather than motivation. Such an interpretation begins to look feasible when we find that nAch also had a near zero correlation with Career Choice ($r = .098$ $p < .05$). Again, on theoretical grounds it was expected that pupils aiming at careers higher up on the socio-economic scale would have relatively high nAch scores too. But this does not appear to be happening. On the other hand we have seen that IQ, DT, English and Arithmetic all have positive correlations ranging from .543 to .414 with Career Choice. Thus, the trend of correlations in the present study suggests that motivation to achieve may be assessed more effectively by asking pupils specific questions related to real life achievement rather than by the analysis of imaginative stories.

This is not to deny that the TAT may still be a valid measure in the area of personality research or in clinical situations. But, as a measure of the extent to which school children may be motivated to do well (or achieve) in academic and non-academic spheres of activity, it appears to be of doubtful value. This is particularly so when the time spent on its administration and scoring is taken into account. Also, when we consider that almost 41% of the protocols in the present study had no achievement imagery in them*, the suggestion made earlier in this chapter that the task of writing stories about these rather ambiguous pictures failed to engage the pupils' interest finds some empirical support. It is possible that the sense of "competition with a standard of excellence" found amongst the American high school and college students tested by McClelland and his colleagues in the later forties and early fifties no longer preoccupies the majority of students nowadays. As works like The Greening of America (Reich 1970) and the Dainton Report (Dainton 1968) point out, there seems to have been an important shift in the values considered to be worthwhile by young people and an outstanding aspect of this change is the conscious playing down of competition. Coleman's(1960) and Williams'(1965) research also suggests that although a sense of competition is not entirely absent from the adolescent sub-cultures they studied, the values held by their subjects are group oriented rather than concerned with individual excellence .

---

*This can be seen from Table VII.2 above where the number of cases under the score category zero is 44 for nAch. Finlayson (1972) has also reported finding a predominance of low scores for nAch with his sample of boys in the academic streams of secondary schools. Therefore, the presence of a number of "Leavers" in the present sample cannot account for the generally low level of nAch found.

Other studies which have specifically looked at the degree of correspondence between questionnaire measures of achievement motivation and TAT measures have generally reported negative findings. McClelland et al (1953) found that nAch was not related to either the Strong Vocational Interest Inventory or the Vernon-Allport Study of Values Scales. Similarly, de Charms et al (1955) also found no significant relationship between direct (questionnaire) and indirect (TAT) measures of achievement motivation. More recently, Argyle and Robinson (1962) have reported a rather mixed finding: nAch and Q-ach (questionnaire achievement) total scores had a correlation of .22 ($p<.01$), but when the total scores were divided into approach and avoidance scores the correlations dropped to zero. From this Argyle and Robinson conclude: "There is little evidence here that the projective and questionnaire measures were measuring the same variable". Finlayson (1972) also reports a "lack of relationship between n/achievement and expressed motivation". The latter was "a reality-based [questionnaire] measure of achievement motivation in which the items described patterns of striving and persistent behaviour in achievement oriented situations".

Perhaps due to this lack of any clear relationship between TAT measures of achievement motivation and questionnaire measures or actual achievement, there is a move towards a theory of motivation which takes into account the role which cognitive control plays in determining what goals people set themselves and how they go about achieving them. Murstein (1963) considers that "the cognitive control approach is closely related to the recent trend in psychoanalysis to pay increasing heed to the ego processes in determining perception rather than to the simple supposed need-state derived from periods of deprivation" (p.69).

In a recent book de Charms (1968) has advocated a very radical reconsideration of the psychology of motivation in such a way that it takes into account the validity of "personal knowledge" or the way in which individuals perceive and interpret their own position in relation to their environment. To illustrate his point, de Charms makes a distinction between the perception of self as an "Origin" or as a "Pawn" to "connote the distinction between forced and free...An origin has a strong feeling of personal causation... and it is a strong motivational force directing future behaviour. A Pawn has a feeling that casual forces beyond his control...determine his behaviour. This consitutes a strong feeling of powerlessness or ineffectiveness" (p.273-274).

In a series of recent papers, Weiner et al (1970, 1971, 1972) have proposed an attributional theory of achievement motivation which seems conceptually to have much in common with the ideas of de Charms described above. According to Weiner et al (1971) "The model is based upon the assumption that beliefs about the causes of success or failure mediate between antecedent stimulus-organism transactions and ensuing achievement behaviour." To an earlier distinction made by Heider (1958) between "can" (ability) and "try" (effort) as determinants of behaviour, Weiner et al add two others of task-difficult and luck. These four elements (ability, effort, task-difficulty and luck) are utilised by individuals to interpret and to predict what has happened/will happen in an achievement situation. Thus, a positive self-perception in terms of ability (i.e. perception of high ability) is likely to be associated with high levels of aspiration or goal-setting for oneself*. The positive correlations between IQ, DT, English and Arithmetic and Career Choice in the present study, in comparison with the non-significant correlations between both the motivation measures (MI and nAch) and Career Choice, suggest that a cognitive control theory of motivation of the type proposed by de Charms (1968) and Weiner et al (1970, 1971, 1972) may be better at explaining these unexpected findings.

Multiple Regression Analysis

The pattern of first order correlations discussed in this chapter so far can be best summarised by presenting the results of a multiple regression analysis carried out to estimate the relative contribution of the independent variables (IQ, DT, MI, nAch, SEB) to the four dependent variables (English, Arithmetic, NAAQ and Career Choice). For each analysis, variables were entered into the equation in a sequence which reflects the value of the first order correlation in descending order.

The table below gives the obtained value of R^2 along with the significance of the regression coefficients:

---

*The term "high" is used in a relative sense here, only implying a comparison with an immediate group such as classmates or other pupils of the same year group at a particular school.

Table VII.3(a) : Multiple Regression Analysis to Estimate the Contribution of Independent Variables to English.

ENGLISH

	R^2	F	df*	p
IQ	.462	26.162	64	***
DT	.524	6.606	63	*
nAch	.535	1.167	62	ns
SEB	.540	0.669	61	ns
MI	.541	0.032	60	ns

Table VII.3(b) : Multiple Regression Analysis to Estimate the Contribution of Independent Variables to Arithmetic.

ARITHMETIC

	R^2	F	df	p
IQ	.400	23.479	64	***
DT	.410	0.525	63	ns
SEB	.423	1.349	62	ns
nAch	.434	1.240	61	ns
MI	.434	0.084	60	ns

* $p < .05$ ** $p < .01$ *** $p < .001$

*Regression coefficients were computed for 66 subjects for whom data were available on all variables.

Table VII.3(c) : Multiple Regression Analysis to Estimate the Contribution of Independent Variables to NAAQ.

	NAAQ			
	R^2	F	df	p
DT	.137	7.235	64	**
SEB	.156	1.386	63	ns
IQ	.157	0.139	62	ns
nAch	.191	1.615	61	ns
MI	.201	0.727	60	ns

Table VII.3(d) : Multiple Regression Analysis to Estimate the Contribution of Independent Variables to Career Choice.

	CAREER CHOICE			
	R^2	F	df	p
IQ	.295	12.464	64	***
DT	.335	3.299	63	ns
SEB	.341	0.738	62	ns
nAch	.344	0.167	61	ns
MI	.344	0.010	60	ns

* $p < .05$ ** $p < .01$ *** $p < .001$

It can be seen from the tables above that the only significant F ratios are for IQ and DT. As would be expected from Wallach and Wing's (1969) argument, IQ does not make a significant contribution to the NAAQ scores, while DT does. Also, while IQ contributes significantly to both English and Arithmetic attainment, the additional contribution of DT is only towards English. Adding on other independent variables beyond DT does not produce a significant rise in the percentage of shared variance (R^2) between SEB, nAch, MI and the four dependent variables. Possible reasons for the low or non-significant correlations between the imaginative story measure of motivation (MI, nAch) and real-life Career Choice have been considered while discussing the first-order correlations. Tables VII.3 (a to d) above highlight the situation regarding these variables. The lack of any further contribution from SEB suggested that the significant correlations of this variable with English, Arithmetic, NAAQ and Career Choice shown in Table 2.1 (Appendix A) may have been due to the variance it shared with IQ. This was confirmed when partial correlations were computed between SEB and the four dependent variables, controlling for the effects of IQ. The partial correlations are given below along with the first order correlations from Table 2.1 (Appendix A) for the purpose of comparison:

Table VII.4 : Partial Correlations Between SEB and Dependent Variables, Controlling for IQ.

	SEB	ENG	ARITH	NAAQ	CAREER CHOICE
Partial	r	.140	.160	.173	.135
	p	$>.05$	$>.05$	$>.05$	$>.05$
First Order	r	.269	.278	.216	.245
	p	$<.01$	$<.01$	$<.05$	$<.01$

The above table shows that the effect of controlling for IQ is least on the SEB-NAAQ partial r. In view of the fact that, of all the independent variables, it is DT that makes a significant contribution to NAAQ in the regression analysis, this is not surprising. Again, the trend of partial correlations confirms that in the present study DT is the most significant predictor of non-academic accomplishments. For example, when DT is controlled for the partial correlations between NAAQ and IQ, SEB, nAch, MI are either non-significant or negative as the following table shows. Again for the purpose of comparison first-order correlations from Table 2.1 (Appendix A) are also included:

Table VII.5 : Partial Correlations Between NAAQ and Independent Variables Controlling for DT.

	NAAQ	IQ	SEB	nAch	MI
Partial	r	.045	.143	- .199	- .159
	p	> .05	> .05	< .05	> .05
First-order	r	.210	.216	.095	- .111
	p	< .05	< .05	> .05	> .05

In the above table it is interesting to note that with the effects of DT removed, the partial correlation between NAAQ and nAch becomes significantly negative and that between NAAQ and MI also has a negative sign, but just fails to reach significance. Here, another measure of real-life achievement (NAAQ) not only fails to be related to the two indices of achievement motivation but it also has a negative relationship with it. This suggests that the TAT measure may indeed be a "fantasy" measure of motivation in the literal sense of the word. In other words, those pupils who actually had an achievement to their credit did not get as high scores for nAch as those who did not have such achievements. Thus, the presence of achievement imagery in the TAT protocols of the latter group may well have been a form of wishful thinking.

CHAPTER VIII

ANALYSIS OF RESULTS FOR THREE LEVELS OF IQ

In Chapter V it was stated that one of the aims of the present study was to test the threshold hypothesis of a decrease in IQ-DT correlations at higher levels of IQ. Another question, related to the threshold hypothesis, was to consider the predictive validity of other variables such as DT, for academic and non-academic achievements at different levels of IQ, and particularly beyond the postulated threshold of IQ 120. Findings regarding these questions will be stated and discussed in this Chapter. But, before this can be done, some detailed comment about how the cut-off points on the IQ scale were arrived at are necessary.

Selection of Cut-off Points to Form High, Middle and Low IQ Groups

The discussion of other studies of the threshold hypothesis in Chapter II showed that one of the main problems encountered in this field is that of non-comparability of IQ-DT correlations at different cut-off points due to the variation in the spread of IQ at these different levels. In the present study this problem was overcome by using a statistical procedure given by T.L. Kelley (1947). On the basis of the known mathematical properties of the unit normal distribution, cut-off points on a scale can be obtained in such a way that each level has the same or a very similar spread of scores. Of course, a necessary requirement of this procedure is that the scores, on the variable being divided up, are normally distributed. Since the distribution of IQ in the present study met this requirement, Kelley's procedure could be used.

Kelley's formulae for working out the mean and standard deviation of a portion and a "slice" of the normal curve are given in Appendix C, along with the computer programme which was written to try out the different possible cut-off points.* A copy of the data print-out is also included in Appendix C to illustrate how the actual cut-off points were chosen. Since a three-level split was required to isolate High, Middle and Low IQ groups, the print-out shows the means and standard deviations for the top portion, the lower portion and the middle portion (or "slice") of the normal curve. The values under the heading "PERCENT" (first left-hand column on the data print-out) stand for the different cut-off points at either end of the normal curve tried out progressively.

As the purpose of using this procedure was to obtain roughly equal standard deviations, we had to consider the lower part of the data print-out to find out where it would be most appropriate to draw the dividing lines. It can be seen from the print-out that when we reach the 20 percent point the standard deviation of the "slice" or middle portion of the distribution comes closest to the standard deviations of the two tails. This cut-off point and the standard deviations have been underlined.

---

*I am very grateful to Dr. A.E.G. Pilliner for suggesting the use of this procedure to obtain comparable High, Middle and Low IQ groups. I am also grateful to my husband, Ali Hasan, for writing the computer programme which saved a considerable amount of my time.

Having determined the cut-off points theoretically in this way, it was necessary to make some adjustment in choosing the actual score levels for forming the High, Middle and Low IQ groups. This had to be done for two reasons: Firstly, the actual distribution of IQ was such that no score actually fell at the 20 percent point at the lower end of the scale. The frequency distribution of IQ is also given in Appendix C after the data print-out, and it shows that the cut-off point could be either at 18.5 percent or at 22.2 percent. Secondly, although there was a convenient cut-off point at the top end at 80.6 percent due to a slight (non-significant) skewness in the data, the theoretically worked out standard deviation (5.267) did not match the actually obtained value of sd at this point (6.087). Therefore, the cut-off points and standard deviations obtained theoretically from Kelley's formulae were used as guidelines only for choosing the actual score levels in such a way as to keep the standard deviations at all three levels as close as possible to the theoretical values. As a result the following cut-off points were chosen:

Table VIII.1 : Cut-off Points on the IQ Scale to Form High, Middle and Low Groups.

IQ Level	IQ Range	n	Cum. Percent	IQ Mean	IQ Stand. Dev.
Low	85 - 102	24	22.2	95.542	5.406
Middle	103 - 122	70	87.0	111.443	5.035
High	123 - 140	14	100.0	129.857	6.075

If we look closely at the reasoning behind the threshold hypothesis in general, as proposed by Anderson, C.E. (1960) it does seem as if Torrance's (1962) argument for extending it to cover the specific question of IQ-DT relationship, is indeed rather tenuous. For example, Anderson clearly stated that "we can think of ability level in terms of thresholds and ask question as to the amount necessary to carry on a task and then consider the factors that determine function beyond this threshold". In other words, Anderson is suggesting that once the basic skills required for doing a task have been acquired the actual level of performance may well depend on other factors. Now, with specific reference to the question of IQ-DT relationship at different levels of IQ, it is difficult to see how it can be assumed that the skills required for doing well on open-ended tests are not fully mastered until an above average level of IQ is reached and then, only beyond that level does performance begin to vary independently of IQ. If it is motor and verbal skill in terms of writing ability that is considered to be the mediating link, then surely there would only be a very small proportion of subjects involved in most studies who would be seriously handicapped in their DT performance as a result of not having sufficient mastery in writing. For we know that most divergent thinking tests (1) do not require any extended piece of writing (it is the ideas that count), (2) do not penalize for bad spelling, writing etc., (3) contain non-verbal as well as verbal sub-tests and (4) are sometimes administered orally. Under these circumstances, if a minimum IQ threshold has to be designated, it is likely to be fairly low down the IQ scale, and certainly not as high as most proponents of the threshold theory have placed it.

From the very mixed (and sometimes contradictory) findings of the studies that have looked at the threshold hypothesis with reference to IQ-DT relationship, it does appear that other variables such as school atmosphere, test instructions and the quality of the test itself also affect the pattern of correlations which emerges. This is very clearly illustrated in Haddon and Lytton's (1968) study when we compare the pattern of correlations they report for the whole sample with the correlations when the analysis is carried out for the formal and informal schools separately. For the whole sample the expected decrease in correlations from the low to high level of IQ is found. But in the separate analyses the correlation dropped to near zero even at the IQ level of "100 and above" in the formal school; while it was .366 for the informal school at the same level of IQ. Clearly it is variables other than IQ level, that seem to be determining the pattern of correlations here. In fact Haddon and Lytton (1968) explain this discrepant finding in terms of the differences in the approach to learning prevalent in the formal and informal schools. They suggest that because of their flexible approach the informal schools maintained "the link between the two aspects [IQ and DT]"; hence the higher correlation at the "100 and above" IQ level. They go on to say that "formal education will tend to destroy this connection by putting a premium on convergent thinking and conformist behaviour". However, at the lowest level of IQ (100 and below) Haddon and Lytton did find the predicted positive correlation in both types of school (formal r = .487, informal r = .548) which lends some support to the threshold hypothesis. But two other recent studies to be discussed below have reported rather contrary findings from which no conclusions regarding the existence of threshold levels can be drawn.

Lytton and Cotton (1969) studied 143 secondary-school children covering a VRQ range of 87 to 137 and obtained the following correlations:

Table VIII.3 : IQ-DT Correlations for Whole Sample and Three Levels of IQ in Lytton and Cotton's (1969) Study.

	n	r
Full range of VRQ	143	.170
VRQ 116+	43	.037
VRQ 101+ (including 116+ category)	114	.141
VRQ 100 and below	29	-.058

The trend of correlations here is rather similar to the one found in the present study and reported in Table VIII.2 above. It should be pointed out that the mean IQ of the whole group in these two studies was also very similar (Lytton and Cotton m = 111 and 112; present study, m = 110), as was the age of the subjects. Therefore, at the secondary stage, factors other than the threshold of IQ seem to be affecting the IQ-DT correlations even more than they did in Haddon and Lytton's (1968) study with primary-school children.

The situation is further complicated when we look at IQ-DT correlations at different levels of IQ reported by Bennett (1973). His study suggests that the sex of subjects may also be one of the "other" variables which influence correlations at different levels, as the following table from his study shows:

Table VIII.4 : IQ-DT Correlations at Three Levels of IQ for Boy and Girls in Bennett's (1973) Study.

Girls	IQ Range	n	r	Boys	IQ Range	n	r
High	115+	53	.38	High	113+	52	.26
Ave.	99-114	59	.31	Ave.	96-112	58	.03
Low	98-	55	.39	Low	95	54	.49

It can be seen that the trend of correlations in this study is contrary to the one reported earlier from Lytton and Cotton's study and the present study. Values of r tend to be lower at the average IQ level, both for girls and boys. Unfortunately, the standard deviation of IQ at the different levels is not reported so the possible effects of score dispersion on these correlations cannot be estimated. Even if restriction of range was not attenuating the correlations at the Average IQ level, there is little evidence, at least in the girls' correlations in favour of the threshold hypothesis. For boys, there is the predicted difference between the correlations at the High and Low IQ levels but the near zero correlation at the Average level does not fit into the pattern.

The negative findings of Cicirelli (1965) and Ginsburg and Whittemore (1968) regarding the threshold hypothesis have been mentioned before (p.37,44). However, in the present study, as in most others (with the exception of Bennett's study discussed above) a low and often non-significant correlation has been found between IQ and DT at the "High" IQ level. With this fact in mind, we can now consider what unique contribution DT makes to academic and non-academic forms of achievement when it is not related to IQ (at the

High and Low IQ levels in the present study) and what the joints contribution of IQ and DT is when they are positively related (at the Middle IQ level).

Correlation Between DT and Dependent Variables at Three Levels of IQ

Table 2.5 in Appendix A shows the correlations of the five independent variables with English, Arithmetic, NAAQ and Career Choice for the High, Middle and Low IQ groups. It is interesting to note in this table that as far as DT is concerned, the only significant correlation it has at the High IQ level is with Career Choice. To the variables of academic (English) and non-academic achievement it is IQ that makes a significant contribution even beyond the postulated threshold. It can be seen from Table VIII.1 in this chapter that the IQ range of the High Group is 123-140 so that if a threshold effect was present it ought to hace shown up in the correlations for this group. On the other hand, at the Low IQ level DT does have a significant correlation with English, Arithmetic and NAAQ. DT is also significantly correlated with English and NAAQ at the Middle IQ level. These correlations at the Middle IQ level cannot be attributed to the positive correlation that has been found between DT and IQ at this level (Table VIII.2, this chapter). It will be noticed from Table 2.5 (Appendix A) that with the exception of English, IQ and DT show a contrasting pattern of correlations at the Middle IQ level. For example, while IQ has significant correlation with Arithmetic and Career Choice, DT has a significant correlation with NAAQ but its correlation with Arithmetic and Career Choice is negligible. Thus, inspite of a significant positive correlation between IQ and DT at the Middle IQ level, these independent variables relate differentially with three out of four dependent variables, showing that even within the IQ range of 102 to

122 (Middle group), IQ and DT are functioning rather independently of each other. Yet at the High IQ level, where the predicted independence of these two variables was expected, it does not show. This pattern of correlations seems to confirm what Burt (1962), Cropley (1966) and Shouksmith (1970) have suggested earlier regarding the complementary nature of convergent and divergent modes of thinking, rather than their complete independence as Getzels and Jackson (1962), Guilford (1950, 1956), Torrance (1962, 1964) and Wallach and Kogan (1965) have argued for. Particularly, the significant positive correlations of DT with academic and non-academic achievement at the Low level of IQ show that where conventionally measured intelligence is not present to any great extent, divergent thinking ability may be of some help in raising the level of performance.

In the present correlations we also have an illustration of what Cicirelli (1965) has called minimum and maximum IQ thresholds. The former refers to the level at which "creativity will begin to distinguish individuals in terms of their academic achievement" and the latter refers to the IQ level beyond which "additional IQ will not distinguish individuals in terms of their academic achievement". On the evidence of the present study it would appear that the minimum threshold is very low indeed, for even within the IQ range of 85 yo 102 (Low IQ group) DT is discriminating between individuals not only in terms of academic achievement but for non-academic accomplishments also. When we look at the IQ-English and IQ-Arithmetic correlations at the three levels of IQ we notice (Table 2.5, Appendix A) that for the High group the correlation between IQ and English barely reaches significance and the IQ-Arithmetic correlation is in fact non-significant. At the same time both these correlations are highly significant at the Middle IQ level. So perhaps the maximum threshold was equally low for the present sample, i.e. beyond IQ 122, additional IQ was not distinguishing individuals in terms of their academic achievement. In Cicirelli's (1965) study it was only at the IQ level of 130-139 that some evidence for a maximum IQ threshold could be found for language achievement only.

Correlations of MI, nAch, SEB with Dependent Variables at Three Levels of IQ

In view of the generally low correlations between the two measures of achievement motivation and other variables in this study, discussed in the previous chapter, it was a matter of some interest, to see if these variables showed any significant associations when three levels of IQ were considered separately. The trend of results for achievement measures is similar to that for DT. That is to say, at the Low IQ level, Motivation Index does correlate significantly with English and Arithmetic attainment. It is noteworthy that these correlations are the only indication of a significant association which MI has shown with any of the variables in the present study, with the exception of a correlation of .331 with nAch. Again, it seems reasonable to conclude that when ability is lacking (as in the Low IQ group - at least relatively speaking) the desire and determination to do well at school may help towards better performance. This conclusion finds further support from the trend of nAch correlations with the attainment measures. These results suggest that perhaps there is a need to extend the threshold hypothesis "downwards" as well, instead of concentrating attention on determining an upper limit only beyond which factors other than ability affect performance. In fact what is called for is a detailed conceptual analysis of the whole idea of a threshold of ability so that more fruitful hypotheses may be formulated regarding the mediating links between ability and performance. It is hoped that the analysis of results in the next chapter (for school "sets" and for the over-lapping cases in the High, Middle and Low IQ groups and the Top, Middle and Leavers' classes) will highlight some of the variables which may have been confounding the relationships predicted by the threshold hypothesis in its present form.

CHAPTER IX

ANALYSIS OF RESULTS FOR SCHOOL CLASSES

The reasons for including a class-wise analysis in this study have already been discussed in Chapter V (p.116-117). In that chapter it was also explained how school classes were combined to form Top, Middle and Leavers' "sets" (p.111-116). Means, standard deviations and the significance of differences between pairs of "set" means are given for all the variables in Table 2.4, (Appendix A). From this table it can be seen that with the exception of the NAAQ means, there is a general downward trend as we move from the Top group to the Leavers' group. For IQ, English, Arithmetic and Career Choice the differences are significant in all three comparisons. For all the variables, with the exception of the two achievement measures (MI and nAch) and NAAQ (where there is no significant difference between the Leavers and the Top set), the Leavers have significantly lower mean scores than the Top and Middle groups. The differences between the Top and Middle groups are not so consistently significant: they differ significantly on IQ, nAch, English, Arithmetic and Career Choice but not on DT, MI, SEB and NAAQ, although the trend of mean scores is in the expected direction. These differences, along with the differences in subject choice, mentioned in Chapter V (p.113-115) suggest that these two groups (Top and Middle "sets") are in fact not very different from what would have been known as A and B streams in a school which formally used streaming.

The note at the foot of Table 1.4 (Appendix A) shows that there is considerable over-lap between these three groups in terms of the range of IQ, so that although the IQ <u>means</u> are significantly different, there are individuals in each group who would not have been there had the classification into High, Middle and Low groups been done only on the basis of IQ, as in the previous chapter. The effects of this over-lap situation will be considered later on in this chapter, but before that is done, a comparison will be made of the way in which independent and dependent variables are related when a correlational analysis is done in terms of class identity rather than levels of IQ. It will of course be remembered that in terms of <u>mean</u> IQ there are levels in this class-wise analysis too, because we have seen (Table 2.4, Appendix A) that the Top, Middle and Leavers' group differ significantly on this variable also.

IQ-DT Correlations for Top, Middle, Leavers' Group

Correlations were also computed between IQ and DT scores for the three groups under consideration here, and the pattern which emerged was very different from the set of IQ-DT correlations reported in the previous chapter (Table VIII.2) for three levels of IQ.

Table IX.1 : IQ-DT Correlations for Top, Middle and Leavers' Groups

	Top	Middle	Leavers
IQ range	101-140	90-131	85-106
IQ-DT r	.457	-.123	.329
p	$< .001$	ns	$< .05$

Comparing the correlations here with those in Table VIII.2 in the previous chapter we notice a complete change in the trend of correlations: for the Top and Leavers' group they are significant and positive; for the Middle group it is negative, but not significant. The change from the Low IQ group to the Leavers' group is not too great and is probably due to the extension in IQ range at the top end (from 102 to 106) and the slight increase in the number of cases (from 24 to 26). For the Top group too, considering the wide range of IQ covered the value of r is only slightly higher than that typically found at this level of IQ (Bennett 1973, Haddon and Lytton 1968). What is most intriguing in the above table is the negative correlation found between IQ and DT for the middle group. Admittedly it is not significant and when comparing it to the Middle IQ level we also have to remember that in terms of the spread of IQ this group is different. But having noted these points it is still difficult to see why there should be such a marked difference in the way IQ and divergent thinking are associated in this group.

Some explanation of the above set of correlations seems possible in terms of the effect that the pupils' sense of belonging to a particular group may have on the way they perform when given a rather unconventional task such as the divergent thinking test. It has been mentioned earlier in this chapter that the three groups under consideration may be indentified in more conventional terms as A, B and C stream pupils. From the sociological point of view it has been argued that there is "a reciprocity of perspective" between A stream pupils and their teachers, "which allows teachers to define, unchallenged by A pupils, ... the nature and boundaries of what is to count as knowledge. It would seem to be the failure of high-ability pupils to question what they are taught in schools that attributes in large measure to their educational achievement" (Keddie 1970). Or, as Katz (1967) has pointed out in his study of low-achieving Negro boys, lack of achievement in this case is due to a discrepancy in the values of the formal education system and those internalised by the low-achieving boys, who in the words of Katz "have internalised a most effective mechanism of self-discouragement....The child, in a sense, has been socialised to impose failure upon himself". With specific reference to the sense of group identity which develops around the classification of pupils into school streams Hargreaves (1967) and Lacey (1970) have shown how this affects academic performance and behaviour in school. These studies also point to the lack of congruence between the values of the lower stream pupils and the school system.

In the context of the present study, although the pupils in the Middle "sets", were still part of the "Certificate" group (unlike the Leavers) they were lower down in the academic hierarchy in comparison with the Top "sets". From their behaviour during the testing sessions, and from some of the answer scripts the impression that came across was that of a clever, but disgruntled group. They were the only group whose members asked questions about why they had to do these tests and if they couldopt out of them. The answer scripts also reflected this attitude. In the Uses sub-test, one of the suggested uses for a sheet of paper was "to write stupid tests like this one on". While doing the tests too, there was a general sense of amusement in this group; whereas the Top and Leavers' groups did all the tests and questionnaires in a rather routine sort of way.

These differences between the Top and Middle sets suggest the possibility that the lack of any significant association between IQ and DT found for the Middle set may be due to the way in which the latter group interpreted the purpose of the open-ended tests. That is to say, the lack of "reciprocity of perspective" between this group and the educational context of the school may have helped them to break away from the notions of appropriateness, correctness etc. which usually apply to academic work, thus taking the new tests in a rather free spirit. Such an approach is likely to minimise the effect of IQ on DT performance; whereas the "academic" approach of the Top group would make them more cautious and self-evaluative even after they had been assured in the test instructions that there were no right answers and it was a test of their imagination. The greater interest of the Middle group in non-academic aspects of school life is also suggested by their slightly higher NAAQ scores in Table 1.4 (Appendix A), although the difference is not statistically significant when these scores are compared with the scores of the Top group.

To ascertain that the difference* in the IQ level of the two groups was not responsible for this contrary trend in correlations, a further analysis of IQ-DT scores was carried out as follows: The overlap in the range of IQ for the two groups (Top, IQ range 101-140; Middle, IQ range: 90-131) made it possible to identify 25 pairs of individuals having the same IQ, one in each group. It was reasoned that if it is the level of IQ that is affecting the IQ-DT correlations in Table IX.1 (as would be expected according to the threshold hypothesis) then there should be little or no difference in the IQ-DT correlations computed for the matched subgroups

---

*Perhaps it ought to be repeated that inspite of the overlap in individual IQs between the two groups, the means differed significantly and the range was different.

from the Top and Middle "sets". In other words, it could be said that the "discrepant" cases in the two sets, i.e. those with different IQs causing the significant difference in the group means (Table 1.4, Appendix A) were also responsible for the difference in the value of correlations for the Top and Middle groups. The following table shows that this was not the case.

Table IX.2 : IQ, DT Means, Standard Deviations, Correlations for Subgroups from Top and Middle Sets, Matched for IQ.

		Top Set	Middle Set
	n	25	25
IQ	mean	112.880	112.800
IQ	sd	7.231	6.983
DT	mean	110.180	111.840
DT	sd	9.501	7.525
IQ-DT	r	.373	-.157
	p	< .03	> .05

It can be seen from the above table that the trend of correlations reported earlier in Table IX.1 continues here, inspite of the fact that individuals in the two groups have been matched for IQ. Clearly then, it is some variable or variables other than IQ which are responsible for the difference in the correlations. In the present study these have not been identified in any specific way, but the suggestion has been made that pupils' sense of identity within the school and their interpretation of what is required of them in a particular testing situation may be two of the factors involved. The present analysis has also shown that to talk of IQ thresholds without taking into account these other factors is an over-simplification.

Correlations Between Independent and Dependent Variables for Top, Middle and Leavers' sets

If we turn to Table 2.4 (Appendix A) again, we notice that as far as the two cognitive measures (IQ and DT) are concerned, they have hardly any predictive validity for the Middle group. With the exception of a low positive correlation between IQ and Arithmetic and a low negative correlation between DT and Arithmetic, both of which just reach significance at the 5% level, neither IQ nor DT has any significant association with English, NAAQ and Career Choice for this group. On the other hand the correlations for the Top group show that DT has a significant positive relationship with all four dependent variables, suggesting that DT does have some predictive validity, not only for concurrent achievement (academic and non-academic), but also for the future aspirations of these pupils. The fact that IQ has a significant correlation only with English, further reveals the complementary nature of measured intelligence and divergent thinking as modes of cognitive functioning. That is to say, where IQ does not succeed in predicting actual performance, DT does. But this interpretation can only be made with an important qualification: IQ and DT are complementary only when they are positively associated with each other, as in the case of the Top group here.

On the face of it, the above statement may sound too self-evident to be worth making, but if we look closely at the pattern of correlations obtained for the Top and Middle groups here, the extent to which IQ and DT are themselves related or not related seems to be crucial in determining the predictive validity of these variables for the criteria of achievement and aspiration under consideration in this study. It will be recalled that some recent studies which have attempted to show that real-life accomplishments of a non-academic kind can be predicted from DT scores, have reported a low or non-significant association between IQ and DT*. From this it has been argued that IQ and DT make independent contributions to different criteria of achievement (Wallach and Wing 1969). However, it has been pointed out earlier (p.76,80) that due to a possible volunteer bias and the mode of analysis used (comparison of top third with lower third students, missing out the middle group) Wallach and Wing's findings cannot be considered conclusive. Similarly in Kogan and Pankove's (1972) study, the only significant, long-term contribution even towards non-academic accomplishments was that of IQ (p. 83).

It is being suggested here that the very lack of a significant association between IQ and DT, which Wallach (1970) considers "paradigmatic" for the validity of divergent-thinking tests, may to some extent be responsible for the inconclusive findings regarding the predictive validity of DT, in studies which have reported low IQ-DT correlations. There is some evidence to support this argument in the predictive validity studies discussed in Chapter III, where academic attainment has usually been employed as a criterion of achievement. For example, as the chart in Appendix D shows Flescher (1963) reports an average IQ-DT r of .09 and, yet he found no evidence for an independent contribution of DT to school achievement. In Cicirelli's (1965) study also, IQ-DT correlations were low (r ranged from .09 to .24, see Appendix D) and invariably IQ had higher correlations with measures of attainment than DT had.

---

*See for example the value of average correlations worked from Kogan and Pankove's (1972) data given on p. 19. Wallach and Wing (1969) also reported that IQ-DT correlations for their sample ranged from -.07 to +.09 only.

On the other hand, studies like those of Cline et al (1962, 1963) or Getzels and Jackson's (1962), where there was a positive association between IQ and DT, the contribution of DT to achievement was more evident.

Going back to the data from the present study, Table 2.4 (Appendix A), shows that, contrary to the trend of more significant correlations between DT and the dependent variables found for the Top group, IQ seems to be a better predictor of academic achievement, and DT of non-academic accomplishments for the Leavers, again suggesting the complementary nature of the contributions. As was the case with the Low IQ group, for the Leavers also, Motivation Index is positively associated with school achievement. The two significant negative correlations for the Middle group in Table 2.4 (Appendix A) call for some comment. In view of the arguments presented earlier in this chapter regarding the effects which group-identification and self-perception might be having on the performance of this group, these negative correlations further suggest that peer-group pressures and overall attitude to school may be more influential in determining the performance of this group rather than measures of cognitive ability or social background which have traditionally been regarded as being important in the past.

CHAPTER X

ANALYSIS OF RESULTS FOR EXPERIMENTAL GROUPS

The discussion of results so far has suggested the complementary status of IQ and DT as modes of thinking rather than their independence from each other. Earlier, in Chapter V, it was mentioned that one way of studying the independent or joint contribution of IQ and DT to measures of academic and non-academic accomplishment and motivation, was to compare the performance of four experimental groups isolated on the basis of their relative standing on IQ and DT. The selection of these groups has been described in Chapter V (p. 120) and results of this analysis will be discussed in the present chapter.

Table 1.6 (Appendix A) gives the basic information regarding the means and standard deviations of all variables for the four experimental groups and for the rest of the sample. From this table it can be seen that the means of the High-High group tend to be the highest of all the five groups, not only for IQ and DT, the two variables on which these groups were selected, but also forother variables. That the IQ and DT means of the High-High group would be the highest was to be expected in view of the procedure used for forming the groups, but the fact that the other means also show the same trend further suggests that the joint contribution of IQ and DT is greater than that of either variable alone. If the analysis of variance to be reported below shows that the differences in the mean scores are significant beyond chance then this would provide additional evidence for the complementary nature of IQ-DT relationship indicated in the data analysis presented so far.

Analysis of Variance Results

Although Table 1.6 (Appendix A) gives the means and standard deviations for the rest of the sample as well, only the four experimental groups were included in the analysis of variance to be presented here. This was done because these are the groups which occupy a truly contrasting position on the two criterion variables, IQ and DT. Any significant differences found in the scores of these groups on the motivation and attainment measures are more likely to highlight the association that IQ and DT may have independently or jointly with these variables. The following table gives the results of the analysis of variance carried out with these four experimental groups:

Table X.1 : Analysis of Variance for Significance of Differences Among Means Scores of Four Experimental Groups.

	F	df	p
MI	0.386	3,64	ns
nAch	4.346	3,64	**
SEB	0.901	3,51*	ns
English	20.335	3,64	***
Arithmetic	24.136	3,61	***
NAAQ	2.085	3,64	ns
Career Choice	7.758	3,48	***

** $p < .01$ *** $p < .001$

---

*No. of cases in this analysis varies because on SEB, Arithmetic and Career Choice complete information was not available for all subjects.

Reference to Table 1.6 (Appendix A) shows that on the four variables with a significant F ratio in the above table, the High-High group has the highest mean scores. It will be noticed that two of these variables (nAch and Career Choice) are measures of aspiration rather than attainment and on both of these also the High-High group comes out significantly better than the other three. Thus, this high ability group is not only performing well in school but in terms of Career Choice also it has set its sights higher than the rest.

In view of the higher correlation between DT and NAAQ than between IQ and NAAQ, shown in Table 2.1 (Appendix A) and discussed in Chapter VII (p.183), it is not surprising to find that the NAAQ mean of the High-DT group in the present analysis also is higher than that of the High-IQ group (Table 1.6 Appendix A). Although the F ratio for NAAQ in the overall analysis just fails to reach significance, it is interesting to note that the NAAQ mean of the High-IQ group is closer to the NAAQ mean of the Low-Low group than to that of the High-High group or the High-DT group. This suggests that when it comes to non-academic achievement, high IQ alone is not accompanied by a high score on NAAQ, whereas those who are amongst the top third of the total distribution on DT but <u>not</u> on IQ show that their achievement in the non-academic field is equal to that of the High-High group. It can be seen from Table 1.6 (Appendix A) that the mean IQ of the High-DT group is only 106.353, much lower than that of the High-IQ group and almost the same as that for the rest of the sample, yet their NAAQ score is highest of all. However, as mentioned earlier in this paragraph, the analysis of variance produced an F ratio for NAAQ which failed to reach significance at the

5% level. Therefore the differences in the NAAQ scores of the four experimental groups discussed here can only be considered suggestive, not conclusive; although the trend of NAAQ means in the present analysis is in the direction that would be expected according to Wallach and Wing's (1969) thesis that DT is a better predictor of NAAQ than IQ.

For the four variables showing significant F ratios in Table X.1, it is possible to draw some conclusions regarding the relative contribution of IQ and DT to measures of academic attainment and motivation with a little more confidence than could be done for NAAQ. McNemar (1962) has pointed out that when a significant F ratio has been obtained in an analysis of variance, then, on the basis of a priori hypotheses, it is legitimate to compare pairs of group means by using the t test. Since a Low-Low group has been included in the present analysis of variance it is possible that this group which has the lowest mean scores on all the four variables (nAch, English, Arithmetic and Career Choice) may be largely responsible for the significant F ratios reported in Table X.1. If this were the case then it could be argued that the superiority of the High-High group suggested by the trend of mean scores in Table 1.6 (Appendix A) may not be statistically significant when scores of the Low-Low group are excluded from the analysis of variance.

There were two ways of checking on the above-mentioned possibility. Firstly, an analysis of variance could be carried out with only the High-High, High-IQ and High-DT groups to see if the significant F ratios dwindle to non-significance as a result of removing the Low-Low group. Secondly according to McNemar's (1962) suggestion, the means of the High-High group could be compared with those of the High-IQ and High-DT groups in turn, to find out if even amongst these three groups, the High-High group does have mean scores that are significantly higher than those of the other two groups. Applying the first of these checks (i.e. analysis of variance with the Low-Low group excluded) produced the following results:

Table X.2 : Analysis of Variance for Significance of Differences Among Mean Scores of Three Experimental Groups.

	F	df	p
MI	0.402	2,48	ns
nAch	2.730	2,48	ns
SEB	0.992	2,39	ns
English	4.418	2,48	*
Arithmetic	9.905	2,45	***
NAAQ	0.916	2,48	ns
Career Choice	6.141	2,37	**

* $p < .05$ ** $p < .01$ *** $p < .001$

Comparing the significance of F ratios in the above table with those in Table X.1, it can be seen that the differences in nAch scores for the three groups are no longer significant, but those for English, Arithmetic and Career Choice remain significant. Inspite of these significant F ratios, no definite conclusions could be drawn regarding the overall superiority of the High-High group unless it could be shown in paired comparisons that their performance was significantly better than the performance of the other two groups. The results of these comparisons are presented below.

Paired Comparisons of Three Experimental Groups

The three variables for which paired comparisons are legitimate (i.e. which have shown a significant difference in the overall analysis) are English, Arithmetic and Career Choice. Results of these comparisons are shown in the following table. Mean scores from Table 1.6 (Appendix A) are also given for ease of comparison.

Table X.3 : Means and Significance of Differences Between Pairs of Means of Three Experimental Groups.

		High-High	H-IQ	H-DT
English	m	121.123	115.306	111.747
	H-H	-		
	H-IQ	ns	-	
	H-DT	*	ns	-
Arithmetic	m	119.250	116.900	107.147
	H-H	-		
	H-IQ	ns	-	
	H-DT	***	***	-
Career Choice	m	5.438	4.500	4.417
	H-H	-		
	H-IQ	**	-	
	H-DT	**	ns	-

* $p<.05$ ** $p<.01$ *** $p<.001$

From the above table it can be seen that the only variable for which the High-High groups shows a significantly higher score than the two other groups is Career Choice. On the two attainment measures the High-High and High-IQ groups do not differ significantly from each other, which suggests that the additional contribution of DT to the attainment scores of the High-High group is not large enough to make a statistically significant difference, although it does tend to pull up the means of this group. There is however, an important difference in the trend of the English and Arithmetic mean scores which must be noted. For English, the means of the High-IQ and High-DT group also do not differ significantly; for Arithmetic they do. And, when we look at the actual value of mean scores for English, the difference between the High-IQ and High-DT groups is much smaller than that between either of these groups and the High-High group. Since the number of subjects in each of these groups was the same for English (n=17) it was possible to carry out orthogonal comparisons by combining the scores of the High-IQ and High-DT group against the scores of the High-High group. The results of these comparisons are set out in the table below.

Table X.4 : Analysis of Variance for Orthogonal Comparisons of English Scores of Three Experimental Groups.

Groups	F	df	p
H-IQ versus H-DT	1.247	1,48	ns
H-IQ + H-DT vs. H-H	7.568	1,48	$<.01$

From the above analysis the superiority of the High-High group over the other two (High-IQ and High-DT) emerges more clearly than it did from the t-test comparisons presented in Table X.3 above. Thus, according to this analysis the joint contribution of IQ and DT to English attainment is reflected in the significantly higher mean score of the High-High group. But, as Table X.3 shows the situation with Arithmetic attainment is rather different. Here the difference between the means of the High-High and High-IQ groups is small when compared with the difference between the mean of the High-DT group and the other two. Considering the comparatively lower correlation between Arithmetic and DT shown in Table 2.1 (Appendix A) it is not surprising that the additional contribution of DT to the Arithmetic attainment of the High-High group is negligible when compared with the attainment of the High-IQ group. Also, as Table 1.6 (Appendix A) shows the mean Arithmetic scores of the High-DT group are slightly less than the scores of the rest of the sample - a group which is high neither on IQ nor on DT. This suggests that a "high" standing on DT alone is associated with rather poor performance in Arithmetic. Again, this was to be expected in view of the purely verbal nature of the open-ended tests on which DT scores are based in the present study. Unfortunately, orthogonal comparisons of the type made for English could not be made for Arithmetic to show more conclusively that the scores of the High-DT group were indeed significantly lower than the combined scores of High-High and High-IQ groups because of unequal numbers in the groups.

The t test comparisons for Career Choice in Table X.4 show that the High-High group had set themselves comparatively higher goals for the future than the other two groups. There is another incidental piece of information in regard to Career Choice which ought to be mentioned here. Table 1.6 (Appendix A) shows that only one pupil in this group had not given a Career Choice, whereas there were five such cases in each of the other two groups. In the light of the cognitive theory of motivation (de Charms 1968, Weiner et al 1970, 1971, 1972) discussed in Chapter VII (p.200) this difference between the groups may be interpreted as reflecting a greater sense of personal control and self-direction on the part of the high ability group when compared with the High-IQ and High-DT groups.

Thus, on the basis of the present analysis, the evidence in favour of the hypothesis that the joint contribution of IQ and DT to attainment and aspiration is greater than that of either IQ or DT alone, is clear only for English attainment and Career Choice. For Arithmetic attainment and NAAQ the contributions of IQ and DT seem to be complementary rather than additive. For example, in the case of Arithmetic attainment the High-IQ group shows performance which is equal to that of the High-High group; for NAAQ it is the High-DT group which compares equally with the High-High group. It will be recalled that the multiple regression analysis reported at the end of Chapter VII also showed the same pattern of relationships for the whole sample: IQ was the only variable that made a significant contribution to Arithmetic (Table VII.3b, p.203) whereas for NAAQ it was DT that had a significant regression coefficient (Table VII.3c, p.204).

Comparison of Experimental Groups for Participation in Extra-Curricular Activities

In Chapter V it was mentioned that besides the administration of tests and questionnaires some further information regarding the extra-curricular activities and hobbies of pupils was also available from official records. This information was utilised to make further comparisons between the experimental groups to see if the higher NAAQ scores of the High-High and High-DT groups were also accompanied by a record of greater participation in activities of a non-academic nature. Information on this subject was available from two sources. The first of these is the school record-card which the school maintains for each pupil from the time of his starting secondary education till the time of leaving school. Besides being a record of educational progress in terms of school examination marks, teachers are also expected to note down in these cards their own comments regarding the pupil's behaviour in class, his/her participation in activities outside the classroom and special home circumstances which come to their notice. Obviously, information of this kind is not very reliable as there is no systematic procedure laid down for recording, and teachers vary in how often and in what detail they record their comments. Therefore, information from this source must necessarily be considered with caution.

The second source of information regarding participation in extra-curricular activities was a questionnaire which pupils fill out for the records of the Careers Office of the local education authority. Since this is a self-report questionnaire rather like the NAAQ, it was reasoned that a comparison of the three experimental groups on their answers to it would throw some light on the validity of the NAAQ as a measure of pupils' participation in non-academic activities. In other words, if NAAQ is a valid index of extra-curricular activities then the trend of NAAQ scores of the three experimental groups should also be reflected in the Careers Office questionnaire in the form of more participation mentioned by the High-High and High-DT groups in comparison with the High-IQ group. For the purpose of analysis the number of activities mentioned by pupils was divided up into three broad categories - none, 1 to 3 and 4+. The following table shows how the three experimental groups reported their own level of participation.

Table X.5 : Number of Pupils in Three Experimental Groups Reporting Participation in Extra-curricular Activities.

	High-High n = 17	High-IQ n=17	High-DT n=17
No. Participation	0	10	4
1 to 3 Activities	9	1	7
4+ Activities	8	6	6

$\chi^2 = 17.375$, $p < .01$ for 4df.

From the above table it can be seen that the two groups with a higher NAAQ score (High-High and High-DT) also reported greater participation in extra-curricular activities on another occasion also, which suggests that although the differences in the NAAQ scroes were not statistically significant they were reflecting a trend which may well have some validity.

Teachers' comments regarding participation in extra-curricular activities were also analysed in the same way as that reported above. The following table presents the results of this analysis.

Table X.6 : Number of Pupils in Three Experimental Groups Reported by Teachers to have Participated in Extra-Curricular Activities.

	High-High n=17	High-IQ n=17	High-DT n=17
No. Participation	6	12	5
1 to 3 Activities	3	4	6
4+ Activities	8	1	6

$\chi^2 = 10.016$, $p < .05$ for 4df.

Again, there is a significant trend in the above table for a greater number of pupils from the High-High and High-DT groups to be reported by teachers as participating in extra-curricular activities. In an earlier study Hasan (1965) has suggested that the high ratings which teachers gave to High Creativity and Low IQ pupils may be reflecting teachers' recognition of these pupils' achievement in non-academic fields of activity, although at the same time the teachers preferred pupils who were performing better academically. The figures in Table X.6 above provide some support for this view as we notice that the trend in this table is in the same direction as that found in the self-report measures (i.e. the Carrers Office questionnaire and the NAAQ). Thus, inspite of the rather unsystematic nature of the records from which information for the present analysis was derived, it does appear that teachers were aware of the out-of-class activities of pupils in our three experimental groups, and there was a reasonable degree of agreement between teacher reports and self-reports. However, comparing Table X.5 and Table X.6 we do find that the greatest discrepancy between the two tables is in the "No Participation" category in which teachers' records have placed more people than the self-reports. This is not surprising since there is no systematic procedure according to which teachers collect information for these records and probably the more outstanding cases come to their notice; whereas in the self-report questionnaire pupils may have mentioned even participation of a routine type which is not likely to have been noted by the teachers.

Admittedly, school records and self-reports of extra-curricular activities would have to be considerably refined and improved before they can be utilised as a reliable source of information regarding pupils' non-academic achievements. For example, some kind of qualitative check would also be desirable in addition to the purely qualitative one used here. Nevertheless, the present analysis does suggest that if the basis of prediction in education is to be broadened to include divergent thinking, one of the criteria which is likely to predict either singly or jointly with IQ, is that of non-academic achievement. In other words, as Holland et al (1962, 1964) and Wallach and Wing (1969) have pointed out, the definition of achievement would have to be extended beyond the traditional academic activities which so far have been regarded as being "extra-curricular".

CHAPTER XI

SUMMARY AND CONCLUSIONS

The main issue under consideration in this study has been that relating to the distinction made by Guilford (1950) between convergent and divergent thinking as modes of cognitive functioning. A review of Guilford's paper showed that although his main discussion was about intellectual abilities he also indicated in the same paper that these abilities cannot be considered in isolation from motivational and other aspects of personality. Thus, convergent and divergent modes of thinking are to be conceived of as two rather pervasive cognitive styles which characterise an individual's behaviour and performance in everyday life.

Since the publication of Guilford's paper, workers in this field (including Guilford and his colleagues) have been attempting to substantiate the grand theory he has presented in the "structure of intellect" model. But the way forward has been blocked by two main obstacles. Firstly, there has been the question of whether the concept of divergent thinking, as measured by performance on open-ended paper - and - pencil tests, can be equated with the concept of creativity in the sense of distinguished and original achievement. Secondly, there is the related question of the extent to which divergent thinking is in fact different from or related to conventional measures of intelligence.

Some studies, usually with a psycho-dynamic approach (Roe, 1952, 1953, Mackinnon, 1962b), have considered creativity only in terms of outstanding performance and they are not of direct relevance to the issue of divergent thinking as measured by paper - and - pencil tests. At the same time, a considerable amount of research has accumulated around the open-ended tests devised by Guilford and his colleagues as measures of divergent thinking. A major concern of these studies (Flescher 1963, Getzels and Jackson 1962, Haddon and Lytton 1971, Kogan and Pankove 1972, Wallach and Wing 1969) has been to consider the validity of the new tests for predicting academic and non-academic criteria of achievement. Thus, although the term "creativity" has been used in American research to indicate a high level of performance on open-ended tests, these studies have little in common with those other studies in which creativity is defined as an original and outstanding contribution towards art, literature, science etc.

The validity studies reviewed in the Section "Background of the Present Study" (pages 11 to 107) deal with two related questions: the convergent and discriminant validity of divergent thinking tests and the predictive validity of these tests. Wallach and Kogan's (1965) conclusion regarding the first of these questions was largely negative : that is they found little evidence in support of the concept of a single dimension of divergent thinking comparable to the concept of general intelligence as measured by conventional intelligence tests. These negative findings Wallach and Kogan attributed to the inappropriateness of a formal and timed testing procedure for the divergent thinking tests used in most of the earlier studies. On the basis of an associative model of creativity proposed by Mednick (1962), they also suggested that divergent thinking tests would have more validity as the measure of "a dimension of individual differences that is as unitary and cohesive as that of general intelligence", if they are scored for associational fluency (number of ideas) and originality (uniqueness of ideas). In a later review, Wallach (1970) also concluded that "the concept of ideational fluency may be paradigmatic for the kind of cognitive performance that is maximally cohesive in itself and maximally distinguishable from convergent thinking".

However, a consideration of studies which have been reported since Wallach and Kogan's (1965) and Wallach's (1970) reviews, showed that an informal testing context is not always associated with higher or in any sense more valid divergent thinking scores (see for example studies by Kogan and Morgan 1969; Kogan and Pankove 1972; Leith 1972; Vernon 1971; Ward, Kogan and Pankove 1972; examined in Chapters II and III). As for the second proposal made by Wallach and Kogan (1965) and Wallach (1970) regarding the scoring procedure, the average correlations reported in Wallach's review and in the present thesis (Chapter II) do tend to be higher among divergent thinking tests and lower between these tests and IQ when scoring is done for number of ideas; whereas there is little difference in the values of average correlations when divergent thinking tests are scored for flexibility or originality alone.

But, it has been argued in this thesis that to obtain the convergent and discriminant validity of divergent thinking tests on the basis of a particular scoring system leaves the question of the construct and predictive validity of these tests still in doubt. For example, in the present study it was found that pupils at the lower end of the IQ scale tended to produce a rather repetitious string of ideas in contrast to those at the higher end who produced a more varied and ingenious set of ideas (p.157). Scored for fluency, the former would get a higher score, whereas flexibility scoring would be to the advantage of the latter. Hence the lower IQ-DT correlations in studies scoring for number of ideas. It was also pointed out in Chapter III (p.48-61) that most studies reporting low IQ-DT correlations also failed to find any evidence for the predictive validity of divergent thinking tests, so that obtaining the convergent and discriminant validity of these tests appears to be a rather academic exercise.

One way of accounting for the poor predictive validity of divergent thinking tests has been to invoke the threshold hypothesis of the decreasing influence of intellectual ability as we go up the IQ scale and the increasing influence of motivational and other factors in determining performance. According to this view, at higher levels of IQ (usually set around 120) divergent thinking begins to function independently of IQ and it is only at this point that we can begin to talk of the validity of divergent thinking tests for predicting performance. However, the review of literature in Chapter II (p.33) shows that the evidence for this view is hardly conclusive when the criterion of performance is academic achievement.

Another line of reasoning employed to account for the low predictive validity of divergent thinking tests is that the criteria of achievement usually employed (i.e. academic grades/examination marks) are inappropriate when divergent thinking is used as a predictor. The alternative suggested is to include non-academic achievements or activities which pupils undertake out of their own initiative and interest (Wallach and Wing 1969). Thus, the importance of motivation in determining performance is also brought into consideration, although Wallach and Wing do not examine the influence of motivational factors in any specific way.

The present study was undertaken to examine three questions arising out of the discussion summarised above. These were:

(a) Is there evidence of a decreasing relationship between IQ and divergent thinking as we move up the IQ scale?

(b) What is the relative contribution of IQ and DT to academic and non-academic achievement, for a wide range of IQ, and at different cut-off points on the IQ scale?

(c) Is there a significant relationship between achievement motivation and performance in the academic and non-academic spheres, especially beyond the IQ level at which performance begins to vary independently of IQ?

For the purpose of analysis pupils were grouped together in three different ways. Firstly, to trace the pattern of IQ-DT correlations at different levels of IQ, a statistical procedure was used to form groups in such a way that the spread of IQ at all three levels was very similar. An examination of the studies of the threshold hypothesis in Chapter II had shown that one of the reasons why findings on this question were inconclusive was that correlations for different levels of IQ were not comparable due to heterogeniety of variance. It was therefore considered necessary to use a procedure which dealt with this problem before going on to a correlational analysis of results. Secondly, to examine the effects of membership of school classes or "sets" on level of performance and on correlations between the different variables under consideration, a further analysis was carried out with the Top, Middle and Leavers' sets. And finally four experimental groups (High IQ - High DT, High IQ - Low DT) were formed on the basis of the individuals' relative standing on IQ and DT, to see what the differential or joint contribution of these variables is to measures of achievement and aspiration.

The analysis of results presented in Section Three may be summarised as follows:

1. Inspite of a reasonable degree of convergent and discriminant validity found for the divergent thinking tests used in this study, there was a moderate and highly significant correlation between IQ and DT over a wide range of IQ (range 85 to 140, r = .467).

2. IQ was also positively associated with measures of academic attainment and future aspiration in terms of the career choice given by pupils.

3. When a questionnaire measure of non-academic accomplishments was used as the criterion of achievement, IQ and DT showed a differential pattern of association with it. In the overall analysis, as well as in the analysis with subgroups, it was DT which showed a consistent trend of moderate positive correlations with NAAQ, whereas IQ was more closely associated with academic attainment and career choice. A comparison of the NAAQ scores of the three experimental groups (High-High, High-IQ and High-DT) with self-reports of extra-curricular activities given in the Careers Office questionnaire and with teachers' records of such activities, suggests that the NAAQ is a valid measure of out-of-class activities and may be a useful research instrument if the criterion of achievement is to be extended to include these activities.

4. In view of 2 and 3 above, the relative contributions of IQ andDDT to the criteria of attainment and aspiration used here, seem to be complementary rather than independent or additive.

5. In regard to the threshold hypothesis, the results of the present study suggest a strong interaction between situational factors (such as school class membership) and ability, because two groups of pupils matched for IQ but belonging to different school classes showed a rather contrary pattern of IQ-DT correlations. Also, not only was there a lack of any significant association between IQ and DT at the high IQ level (as would be predicted by the threshold theory) but at the low IQ level too IQ and DT did not show a significant correlation. Therefore, it seems rather an oversimplification to consider level of IQ alone as the crucial variable which determines IQ-DT correlations. Data from the present study would produce an inverted u curve to represent the pattern of IQ-DT correlations (low-high-low for three levels of IQ) rather than a decreasing value of r as we move up the IQ scale, which is the pattern that ought to be obtained according to the ability gradient theory.

6. The two measures of achievement motivation (MI and nAch) had a significant positive correlation with each other ($r = .331$, $p < .001$) although they were derived from two rather different scoring schemes. In the present study they proved to be of little value as predictors of either academic attainment or non-academic achievement, or career choice when results were analysed for the whole sample, as the regression analysis reported in Chapter VII shows. However, when the pupils were grouped according to levels of IQ or membership of Top, Middle or Leavers' "sets" some significant relationships did appear. For example, for the Low IQ group (range 85-102) both MI and nAch showed a significant correlation with measures of academic attainment. Similarly, for the Leavers' "set" MI was significantly correlated with English and Arithmetic. This trend of results suggests that contrary to the prediction of the threshold hypothesis that

beyond a certain level of ability, motivational factors begin to make a significant contribution to performance, it is at the _below_ average level that the measures of motivation used here seem to be of most importance.

In this study an attempt has been made to explore an area in which evidence regarding some basic issues, such as the relationship between IQ and divergent thinking, and the validity and reliability of divergent thinking tests is confused, and even contradictory. One author in the field has remarked that the psychological literature on creativity is like a collection of reports by several blind men on the same elephant (Yamamoto 1965c). If the ever increasing number of journal articles and books on the subject are anything to go by, the present state of affairs can hardly be attributed to lack of effort on the part of interested workers. Already, in 1964, George Miller had testily noted that "Creativity is "in". Everyone is talking about it. Educators are trying to recognise it, psychologists are trying to measure it, business men are trying to buy it .."Yet, despite all this effort, the main concepts around which theories and explanations regarding divergent thinking have been built remain unclear. Claims made on the basis of rather equivocal findings - claims which if substantiated have far reaching implications for our ideas about intellectual abilities, their growth and measurement - must still be accepted with caution.

It is hoped that the critical analysis of the literature and the research reported in this study have helped to specify and to clarify some crucial aspects of the creativity-intelligence debate.

Those which in retrospect seem most important are: the effect of scoring procedures and school-class membership on IQ-DT correlations; the question, as yet unresolved, of the validity of the threshold hypothesis to explain a lack of association between IQ and DT at higher IQ levels, when evidence as to that lack is still conflicting; the impact on the threshold hypothesis argument, of the move away from Anderson's (1960) original formulation of the problem, towards a cognitive view of DT as well as IQ; and the importance of motivational factors across the whole IQ range without restriction to levels above an IQ threshold of dubious provenance.

Finally, the present study has also examined the question of the better predictive validity of DT scores in comparison with IQ, for achievements of a non-academic or extra-curricular type. The findings discussed in Chapters VII and X suggest that this is the area in which lies the greatest strength of tests of divergent thinking. In the light of these findings, it is relevant to ask whether the currently accepted criteria of educational achievement should not be broadened to include activities hitherto termed "extra-curricular" - which is tantamount to inviting a fresh appraisal of educational objectives. We cannot believe that tests of divergent thinking tap something useful and at the same time brush off what they test as peripheral. It may well be that such broadening of the definition of educational achievement will have a valuable "pay-off" both for the individual and for the society of which he is a part.

APPENDIX A

Statistical Tables

TABLE 1.1 : DATA FROM (A) FIVE SCHOOL CLASSES (B) WHOLE SAMPLE.
MEANS, STANDARD DEVIATIONS, SIGNIFICANCE OF DIFFERENCE
BETWEEN CLASS MEANS
(Girls and Boys)

VARIABLES	CLASSES	4C1	4C4	3C1	3C4	LEAVERS	TOTAL
IQ	n	26	19	17	20	26	108
	m	115.885	109.053	121.000	111.540	97.731	110.296
	sd	9.808	9.120	8.155	6.065	6.290	11.264
	4C1	-					
	4C4	*	-				
	3C1	ns	***	-			
	3C4	ns	ns	***	-		
	LEAVERS	***	***	***	***	-	
DT	n	26	19	17	20	26	108
	m	113.019	112.868	116.912	110.350	102.712	110.630
	sd	12.279	7.259	8.771	9.903	8.620	10.683
	4C1	-					
	4C4	ns	-				
	3C1	ns	ns	-			
	3C4	ns	ns	*	-		
	LEAVERS	***	***	***	**	-	
MI	n	26	19	17	20	26	108
	m	122.846	100.158	99.059	121.800	106.192	110.907
	sd	29.362	37.952	30.748	21.895	25.379	31.476
	4C1	-					
	4C4	*	-				
	3C1	*	ns	-			
	3C4	ns	*	*	-		
	LEAVERS	*	ns	ns	ns	-	
nAch	n	26	19	17	20	26	108
	m	3.692	2.529	5.529	3.200	1.962	3.269
	sd	3.271	2.875	3.430	3.156	2.749	3.249
	4C1	-					
	4C4	ns	-				
	3C1	ns	**	-			
	3C4	ns	ns	*	-		
	LEAVERS	*	ns	**	ns	-	
SEB	n	23	15	15	17	20	90
	m	4.130	4.133	4.200	4.118	2.850	3.856
	sd	1.058	1.302	1.265	1.219	1.387	1.329
	4C1	-					
	4C4	ns	-				
	3C1	ns	ns	-			
	3C4	ns	ns	ns	-		
	LEAVERS	**	**	**	**	-	
ENG	n	26	19	17	20	26	108
	m	115.869	109.053	121.006	111.450	97.731	110.294
	sd	9.799	9.109	8.039	5.998	6.308	11.242
	4C1	-					
	4C4	*	-				
	3C1	ns	***	-			
	3C4	ns	ns	***	-		
	LEAVERS	***	***	***	***	-	

*p < .05, ** p < .01, *** p < .001

(cont).

Table 1.1 (cont.)

VARIABLES	CLASSES	4C1	4C4	3C1	3C4	LEAVERS	TOTAL
ARITH	n	24	18	16	19	26	103
	m	115.354	109.027	120.750	111.579	97.731	109.942
	sd	9.885	9.265	8.351	6.106	6.192	11.245
	4C1	-					
	4C4	*	-				
	3C1	ns	**	-			
	3C4	ns	ns	**	-		
	LEAVERS	***	***	***	***	-	
NAAQ	n	26	19	17	20	26	108
	m	9.384	11.105	10.412	10.000	7.846	9.611
	sd	5.797	3.557	5.927	4.800	4.173	4.971
	4C1	-					
	4C4	ns	-				
	3C1	ns	ns	-			
	3C4	ns	ns	ns	-		
	LEAVERS	ns	**	ns	ns	-	
CAREER CHOICE	n	18	12	15	13	19	77
	m	4.889	4.500	5.133	4.539	3.579	4.494
	sd	0.900	0.798	0.915	0.660	0.961	1.021
	4C1	-					
	4C4	ns	-				
	3C1	ns	ns	-			
	3C4	ns	ns	ns	-		
	LEAVERS	***	*	***	**	-	

TABLE 1.2 : MEANS, STANDARD DEVIATIONS, SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS OF BOYS AND GIRLS

(A) WHOLE SAMPLE (B) NON-LEAVERS ONLY

		Whole Sample			Non-Leavers Only		
		GIRLS	BOYS	p	GIRLS	BOYS	p
IQ	n	53	55		44	38	
	m	112.585	108.091	*	115.545	112.816	ns
	sd	11.656	10.510		10.192	8.318	
DT	n	53	55		44	38	
	m	114.123	107.264	**	115.966	109.868	**
	sd	10.367	9.960		10.114	9.084	
MI	n	53	55		44	38	
	m	113.830	108.091	ns	115.705	108.579	ns
	sd	31.462	31.520		32.345	34.139	
nAch	n	53	55		44	38	
	m	3.472	3.073	ns	3.750	3.602	ns
	sd	3.279	3.237		3.321	3.317	
SEB	n	45	45		38	32	
	m	3.800	3.911	ns	4.000	4.313	ns
	sd	1.217	1.443		1.200	1.120	
ENG	n	53	55		44	38	
	m	113.977	106.744	**	116.927	111.208	**
	sd	11.309	10.058		9.832	7.889	
ARITH	n	53	55		42	35	
	m	112.294	107.635	*	115.024	112.914	ns
	sd	11.609	10.479		10.720	7.585	
NAAQ	n	53	55		44	38	
	m	10.396	8.855	ns	11.182	9.000	*
	sd	5.104	4.763		5.077	4.921	
CAREER CHOICE	n	41	36		35	23	
	m	4.683	4.278	ns	4.829	4.739	ns
	sd	0.879	1.137		0.822	0.915	

$*p<.05$ $**p<.01$ $***p<.001$

TABLE 1.3 : MEANS, STANDARD DEVIATIONS, SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS OF (A) LEAVERS/NON-LEAVERS (B) THIRD YEAR/FOURTH YEAR NON-LEAVERS

(Girls and Boys)

		LEAVERS	NON-LEAVERS	p	3rd Yr.	4th Yr.	p
IQ	n	26	82		37	45	
	m	97.731	114.281	***	115.838	113.000	ns
	sd	6.290	9.414		8.500	10.016	
DT	n	26	82		37	45	
	m	102.712	113.140	***	113.365	113.267	ns
	sd	8.620	10.068		9.845	10.355	
MI	n	26	82		37	45	
	m	106.192	112.402	ns	111.351	113.267	ns
	sd	25.379	33.175		31.597	34.749	
nAch	n	26	82		37	45	
	m	1.962	3.683	*	4.270	3.200	ns
	sd	2.749	3.299		3.445	3.130	
SEB	n	20	70		32	38	
	m	2.850	4.143	***	4.156	4.132	ns
	sd	1.387	1.171		1.221	1.143	
ENG	n	26	82		37	45	
	m	97.730	114.277	***	115.841	112.991	ns
	sd	6.308	9.379		8.428	10.005	
ARITH	n	26	77		35	42	
	m	97.731	114.069	**	115.771	112.643	ns
	sd	6.192	9.426		8.485	10.022	
NAAQ	n	26	82		37	45	
	m	7.846	10.171	*	10.243	10.111	ns
	sd	4.173	5.094		5.273	5.001	
CAREER CHOICE	n	19	58		28	30	
	m	3.579	4.793	***	4.857	4.733	ns
	sd	0.961	0.853		0.848	0.868	

*p<.05 **p<.01 ***p<.001

TABLE 1.4 : MEANS, STANDARD DEVIATIONS, SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS OF TOP, MIDDLE, LEAVERS' SETS

(Girls and Boys)

SCH. CLASS		4C1 + 3C1	4C4 + 3C4	3B2	SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS			
	SETS	TOP	MIDDLE	LEAVERS		TOP	MIDDLE	LEAVERS
IQ	n	43	39	26	TOP			
	m	117.907	110.534	97.731	MIDDLE	***	-	
	sd	9.434	7.684	6.290	LEAVERS	***	***	-
DT	n	43	39	26	TOP	-		
	m	114.767	111.744	102.923	MIDDLE	ns	-	
	sd	11.082	8.729	8.630	LEAVERS	***	***	-
MI	n	43	39	26	TOP	-		
	m	113.209	111.385	106.192	MIDDLE	ns	-	
	sd	31.945	34.715	25.379	LEAVERS	ns	ns	-
nAch	n	43	39	26	TOP	-		
	m	4.419	2.872	1.962	MIDDLE	*	-	
	sd	3.417	3.002	2.749	LEAVERS	**	ns	-
SEB	n	38	32	20	TOP	-		
	m	4.158	4.125	2.850	MIDDLE	ns	-	
	sd	1.128	1.238	1.387	LEAVERS	***	***	-
ENG	n	43	39	26	TOP	-		
	m	117.900	110.282	97.730	MIDDLE	***	-	
	sd	9.393	7.666	6.308	LEAVERS	***	***	-
ARITH	n	43	39	26	TOP	-		
	m	117.513	110.338	97.730	MIDDLE	***	-	
	sd	9.572	7.801	6.192	LEAVERS	***	***	-
NAAQ	n	43	39	26	TOP	-		
	m	9.791	10.590	7.846	MIDDLE	ns	-	
	sd	5.800	4.216	4.173	LEAVERS	ns	*	-
CAREER	n	33	25	19	TOP	-		
CHOICE	m	5.000	4.520	3.579	MIDDLE	*	-	
	sd	0.901	0.741	0.961	LEAVERS	***	***	-

IQ range of Top set: 101-140, Middle set: 90-131, Leavers set: 85-106

*$p < .05$ **$p < .01$ ***$p < .001$

TABLE 1.5 : MEANS, STANDARD DEVIATIONS, SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS OF HIGH, MIDDLE, LOW IQ GROUPS

(Girls and Boys)

	IQ RANGE	123-140	103-122	85-102	SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS			
	IQ LEVELS	HIGH	MIDDLE	LOW		HIGH	MIDDLE	LOW
IQ	n	14	70	24	HIGH			
	m	129.857	111.443	95.542	MIDDLE	***	-	
	sd	6.075	5.035	5.406	LOW	***	***	-
DT	n	14	70	24	HIGH	-		
	m	118.500	111.293	104.104	MIDDLE	*	-	
	sd	10.314	10.304	8.304	LOW	***	**	-
MI	n	14	70	24	HIGH	-		
	m	117.929	110.329	108.500	MIDDLE	ns	-	
	sd	33.831	32.002	29.201	LOW	ns	ns	-
nAch	n	14	70	24	HIGH	-		
	m	3.929	3.486	2.250	MIDDLE	ns	-	
	sd	3.430	3.309	3.848	LOW	ns	ns	-
SEB	n	14	58	18	HIGH	-		
	m	4.071	4.000	3.222	MIDDLE	ns	-	
	sd	0.997	1.199	1.768	LOW	ns	*	-
ENG	n	14	70	24	HIGH	-		
	m	119.779	112.169	99.292	MIDDLE	**	-	
	sd	10.147	9.527	8.114	LOW	***	***	-
ARITH	n	12	70	24	HIGH	-		
	m	117.918	112.649	98.396	MIDDLE	*	-	
	sd	7.556	9.813	7.968	LOW	***	***	-
NAAQ	n	14	70	24	HIGH	-		
	m	10.571	9.957	8.042	MIDDLE	ns	-	
	sd	4.519	5.380	3.617	LOW	ns	ns	-
CAREER	n	12	49	16	HIGH	-		
CHOICE	m	5.333	4.592	3.563	MIDDLE	*	-	
	sd	0.779	0.888	0.892	LOW	***	***	-

*p<.05 **p<.01 ***p<.001

TABLE 1.6 : MEANS, STANDARD DEVIATIONS OF (A) FOUR EXPERIMENTAL GROUPS (B) REST OF THE SAMPLE

(Girls and Boys)

	EXP. GROUPS	H-H	H-IQ	H-DT	L-L	REST
IQ	n	17	17	17	17	40
	m	124.588	120.824	106.353	97.529	106.850
	sd	8.653	5.457	7.237	5.970	6.455
DT	n	17	17	17	17	40
	m	125.291	109.059	121.029	98.382	105.534
	sd	7.461	4.555	5.959	4.404	5.344
MI	n	17	17	17	17	40
	m	117.235	107.529	115.824	108.529	108.575
	sd	28.908	33.165	39.443	28.970	30.008
nAch	n	17	17	17	17	40
	m	5.412	2.765	3.647	1.529	3.150
	sd	3.411	3.615	3.040	2.787	2.887
SEB	n	16	14	12	13	35
	m	4.375	3.786	4.083	3.615	3.657
	sd	0.885	1.051	1.505	1.850	1.305
ENG	n	17	17	17	17	40
	m	121.125	115.306	111.747	98.412	107.993
	sd	9.785	9.689	8.322	7.175	9.460
ARITH	n	16	17	17	17	38
	m	119.250	116.900	107.147	98.294	109.737
	sd	10.621	8.212	5.364	6.790	11.040
NAAQ	n	17	17	17	17	40
	m	11.647	9.471	11.706	8.059	8.575
	sd	4.999	5.811	5.621	3.561	4.431
CAREER	n	16	12	12	12	25
CHOICE	m	5.438	4.500	4.417	3.750	4.280
	sd	0.629	1.000	0.996	1.138	0.792

NOTES FOR CORRELATION TABLES 2.1 to 2.5

1. Decimal points and + signs have been omitted in all correlation tables.

2. Item-total correlations in Tables 2.2 and 2.3 are between independent scores.

3. Means and standard deviations for NAAQ and DT in Tables 2.2 and 2.3 are for raw scores.

4. For n = 108, correlations of .159, .223 and .289 are significant at .05, .01 and .001 respectively (one-tailed test).

TABLE 2.1 : INTER-CORRELATIONS OF INDEPENDENT AND DEPENDENT VARIABLES

VARIABLES	1	2	3	4	5	6	7	8	9
1. IQ	-								
2. DT	467	-							
3. MI	078	075	-						
4. nAch	205	211	331	-					
5. SEB	249	211	038	073	-				
6. ENG	679	539	105	274	269	-			
7. ARITH	632	383	062	245	278	704	-		
8. NAAQ	210	371	-111	-095	216	204	192	-	
9. CAREER CHOICE	543	414	016	098	245	526	494	184	-

TABLE 2.2 : MEANS, STANDARD DEVIATIONS, INTER-CORRELATIONS OF NAAQ

	m	sd	1	2	3	4	5	6	7	8	9	10
1. ART	0.991	1.759	-									
2. DRAMA	0.898	1.824	003	-								
3. DEBATE	0.426	1.201	192	272	-							
4. WRITING	0.806	1.544	099	146	317	-						
5. MUSIC	1.491	2.125	-009	126	-050	192	-					
6. SCIENCE	0.556	0.998	089	088	058	125	-033	-				
7. SOCIAL SERVICE	1.324	2.018	222	169	182	195	196	114	-			
8. LEADERSHIP	0.751	1.789	163	219	111	077	052	099	251	-		
9. SPORTS	4.471	5.666	-027	-087	016	124	006	122	073	099	-	
10. NAAQ TOTAL	7.130	6.400	201	284	303	282	127	145	402	285	071	-

Note: All correlations are between independent scores

TABLE 2.3 : MEANS, STANDARD DEVIATIONS, INTER-CORRELATIONS OF DT (ITEMS/SUBTEST) IQ, ENGLISH AND ARITHMETIC

	m	sd	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
USES TEST																								
1. Brick (F)	3.167	1.589																						
2. Pen (F)	2.907	1.585	677																					
3. Paperclip (F)	2.990	1.912	698	595																				
4. Toothpick (F)	2.833	1.568	659	584	617																			
5. Sheet of Paper (F)	3.444	1.648	287	238	150	311																		
6. USES (F)	15.333	6.349	778	682	665	714	283																	
7. USES (U)	1.704	1.896	602	544	536	666	303	691																
SIMILARITIES TEST																								
8. Potatoes and Carrots (F)	3.954	1.790	377	325	451	387	140	443	307															
9. Cat and Mouse (F)	3.917	2.337	331	311	337	348	146	386	153	556														
10. Train and Tractor (F)	3.482	1.988	416	400	518	467	216	531	408	605	542													
11. Milk and Meat (F)	2.880	1.527	493	412	522	460	155	537	320	611	629	610												
12. Curtain and Rug (F)	2.417	1.772	487	503	572	429	172	570	421	404	521	579	640											
13. SIMILARITIES (F)	16.648	7.614	511	476	582	512	206	602	388	678	677	709	767	640										
14. SIMILARITIES (U)	1.398	1.734	557	493	576	471	212	607	545	385	430	565	572	541	611									
CONSEQUENCES TEST																								
15. Trees (F)	3.694	2.129	307	315	385	393	098	394	163	359	412	333	414	314	454	291								
16. Sun (F)	1.287	1.529	374	323	340	348	-003	362	259	224	321	277	383	552	429	330	421							
17. Speak (F)	2.315	1.791	445	405	468	438	066	479	300	436	531	532	588	556	652	399	582	571						
18. Fly (F)	2.769	1.888	346	302	375	334	078	377	245	423	419	416	561	490	563	402	543	532	530					
19. CONSEQUENCES (F)	10.605	5.913	452	416	488	471	080	501	294	454	526	485	606	580	652	388	624	600	686	647				
20. CONSEQUENCES (U)	2.167	2.247	505	324	437	400	114	467	378	336	392	408	523	475	522	587	372	472	620	561	621			
21. FLEXIBILITY TOTAL (F)	42.046	16.962	661	602	678	650	220	613	535	609	617	679	753	710	722	655	552	527	711	610	651	626		
22. UNIQUENESS TOTAL (U)	5.269	4.799	676	545	625	621	249	711	505	418	400	556	578	584	618	683	344	438	553	506	566	543	741	
23. IQ	110.296	11.264	223	181	259	289	084	272	255	449	311	429	319	311	450	231	313	294	280	434	412	261	447	306
24. ENGLISH	110.294	11.242	340	242	278	360	135	352	233	436	382	453	371	391	503	307	423	382	387	546	543	381	547	381
25. ARITHMETIC	109.942	11.245	286	219	209	350	-015	282	-187	362	309	-270	264	269	363	156	233	327	257	420	379	259	400	250

F = Flexibility Score

U = Uniqueness Score

TABLE 2.4 : CORRELATION BETWEEN INDEPENDENT AND DEPENDENT VARIABLES FOR TOP, MIDDLE, LEAVERS' SET

(Boys and Girls)

DEPEND VAR / INDEP VAR		ENG			ARITH			NAAQ			CAREER CHOICE		
		TOP	MID	LEAV	TOP	MID	LEAV	TOP	MID	LEAV	TOP	MID	LEAV
IQ	r	509	043	621	199	311	626	204	090	069	255	224	179
	p	***		***		*	***						
DT	r	546	103	304	328	-308	267	414	108	454	415	-249	347
	p	***			*	*		**		**	**		
MI	r	-217	280	457	-024	-084	476	-138	-124	014	060	-114	-013
	p		*	**			**						
nAch	r	031	133	251	091	-114	284	-125	-227	-018	-060	-121	-303
	p												
SEB	r	135	-376	344	004	-118	307	066	226	329	039	075	165
	p		*										

*p <.05 , ** p <.01 , *** p <.001

TABLE 2.5 : CORRELATION BETWEEN INDEPENDENT AND DEPENDENT VARIABLES FOR HIGH, MIDDLE, LOW LEVELS OF IQ

(Boys and Girls)

DEPEND VAR / INDEP VAR		ENG			ARITH			NAAQ			CAREER CHOICE		
		HIGH	MID	LOW	HIGH	MID	LOW	HIGH	MID	LOW	HIGH	MID	LOW
IQ	r	475	518	460	-038	484	390	541	095	094	-111	261	210
	p	*	***	**		***	*	*				*	
DT	r	430	398	513	367	161	383	-054	340	349	621	180	296
	p		***	*			*		***	*	*		
MI	r	-196	055	466	190	-039	376	-398	-115	006	-266	-018	058
	p			**			*						
nAch	r	119	201	363	134	148	358	-141	-170	083	045	-046	-065
	p		*	*			*						
SEB	r	178	035	488	-004	122	355	-198	230	365	303	048	397
	p			*					*				

* p <.05 , ** p <.01 , *** p <.001

APPENDIX B : Tests and Questionnaires

TAT

Name: .... Girl / Boy

School: ....

Date of Birth: ....

To-day's Date: ....

Read the following carefully before starting this test

1. This is a test of your imagination. You will be shown some slides on the screen. After you have seen a picture, write a story about it. There are no right or wrong answers so you can make up any story you like.

2. Do not just describe the picture. Be sure to write a story with a beginning and an end. That is, your story should tell:

   (a) what has led up to the scene shown in the picture
   (b) what is happening now
   (c) what are the thoughts and feelings of the people in the pictures, and
   (d) how it will all turn out.

3. Numbered pages are provided for each of the pictures you will see. The four point which you have to keep in mind while writing the stories, are repeated for you on each page.

4. Begin every story on a new page.

5. Remember that there are no right answers.

6. Please write legibly.

Picture No. 1

(a) what has led up to the scene shown in the picture

(b) what is happening now

(c) what are the thoughts and feelings of the people in the picture, and

(d) how it will all turn out.

OPEN-ENDED TESTS

Name: .... Girl/Boy

School: .... Class: ....

Date of Birth: ....

To-day's date: ....

Read the following before starting this test:

1. Like the stories you wrote earlier, this is also a test of your imagination. There are no predetermined right answers to the questions inside.

2. There are three sets of questions and you will have 40 minutes to answer them. Time will be announced after every 10 minutes.

3. Detailed instructions about what you are to do are given before each set of questions.

4. You may work through the questions in any order you like. That is, you may begin somewhere in the middle, not necessarily at the beginning, and come back to a question later.

5. You may answer in note form provided your meaning is clear.

6. If you are short of space for a particular test, continue on the back page, but be sure to give the heading and the question number (e.g. Uses, 2 or Similarities, 3 etc) before you write the answers.

I. Uses for Objects

Listed below are five objects. Write down as many different uses as you can think of for each object. Write anything that comes to mind, no matter how unusual or strange it may seem:

1. BRICKS:

2. PENCILS:

3. PAPER CLIPS:

4. TOOTHPICKS:

5. SHEET OF PAPER:

II. SIMILARITIES

The two words in each pair below have several features in common. Write down all the ways in which you think the two words in a pair are alike. Do not be afraid to use your imagination.

1. A potato and a carrot:

2. A cat and a mouse:

3. A train and a tractor:

4. Milk and meat:

5. A curtain and a rug:

III. CONSEQUENCES

Given below are four imaginary situations. If these situations were real, many other changes would follow and life as we know it would be very different. Write down the different consequences you think would follow in each of these situations. What would the consequences be if:

1. there were no trees in the world?

2. the sun rose in the north and set in the south?

3. the human race could not speak at all?

4. human beings could fly as the birds do?

NON-ACADEMIC ACCOMPLISHMENTS QUESTIONNAIRE

Name: .... Girl/Boy

School: .... Class: ....

Date of Birth: ....

Today's date ....

Read the following before answering this questionnaire.

1. You will find inside statements describing some possible achievements of students. If you think a statement describes you, please put a circle round the letter (a, b, c, etc.) before that statement.

2. You may circle as many statements as you think apply to you.

3. Remember that this questionnaire is about interests and activities which you have pursued for your own satisfaction, in school or at home. You may have done some of the work in school, with advice and suggestions from your teachers, but the important point is that you were involved in an activity because you liked it.

4. Do not include achievements occuring before you transferred to secondary school.

5. Read all the statements in each section carefully, before you circle any. If you wish to change your answer do so very clearly.

ART AND CRAFT

a. I have created art/craft work such as painting, drawing, sculptures, cartoon, photography, pottery, woodwork etc. (not as part of course work).

b. I had art/craft work exhibited in school or outside school.

c. I entered a school or local art/craft competition.

d. I won a prize or award in a school or local art/craft competition.

e. I entered a regional or national art/craft competition.

f. I won a prize or award in a regional or national art competition.

DRAMA

a. I was a member of a drama group in school.

b. I was a member of a drama group outside school.

c. I played a minor role in the cast or crew of plays sponsored by school, community or religious group.

d. I played a major role in a dramatic production sponsored by school, community or religious group.

e. I received a prize or award, from the school, community or other local group for acting, playwriting or any other phase of dramatic production.

f. I received a regional or national award for acting, playwriting or any other phase of dramatic production.

SPEECH AND DEBATE

a. I was a member of a debating club or society in school.

b. I was a member of a debating club or society outside school.

c. I took part in school or local speech contests and debates.

d. I won a prize in a school of local speech contest or debate.

e. I took part in a regional or national speech contest or debate.

f. I won a prize in a regional or national speech contest or debate.

CREATIVE WRITING

a. I have written original stories, essays, articles, plays or poems (not as part of course work) but have not published them.

b. I have published original writing in school paper or magazine.

c. I have edited or worked on the editorial staff of a school paper or magazine.

d. I have published original writing in a paper or magazine outside school.

e. I have won a school or local literary prize for original writing.

f. I have won a regional or national literary prize for original writing.

MUSIC

a. I played a musical instrument for my own pleasure.

b. I sometimes sang or performed music as a soloist or member of a school or community musical group.

c./

c. I have composed music or song which has been performed in school.

d. I have been a regular member of a musical group, orchestra etc., in the school or community.

e. I have won a prize or award in a musical competition.

f. A public performance has been given of music which I have written or arranged.

SCIENCE

a. I am a member of a science club or reading and discussion group.

b. For my own personal satisfaction, I build model aeroplanes, boats etc; carry out scientific experiments, collect insects, flowers etc; or study nature (i.e. birds, stars etc.)

c. I have entered a school or local scientific competition.

d. I have won a prize in a school or local scientific competition.

e. I have entered a regional or national scientific competition.

f. I have won a prize in a regional or national scientific competition.

SOCIAL SERVICE

a. I am a member of a community or religious social service group such as the Red Cross, Scouts, Guides, Boys' Brigade, Young Volunteers etc.

b. I take an active part in the programmes sponsored by a community or religious social service group.

c. I was elected or appointed to a position of responsibility, committee member, secretary, treasurer, president etc. in a local social service organisation.

d. I was elected or appointed to a position of responsibility in a regional or national social service organisation.

e. I received a local award or prize for outstanding social service or community work.

f. I received a regional or national award for outstanding social service or community work.

LEADERSHIP

a. I was a monitor, class-representative or prefect in school

b. I held an office of responsibility (committee member, secretary, treasurer etc.) in a school club or society.

c. I held an office of responsibility (committee member, secretary, treasurer etc.) in a club or society outside school.

d. I was elected or nominated president/chairman of a club or society in school.

e. I was elected or nominated president/chairman of a club or society outside school.

f. I won a prize or award for leadership.

SPORTS AND GAMES

a. I have played as a member of a team in a sports or games competition in school.

b. I have played as a member of a team in a sports or games competition outside school.

c. I have led a sports or games team as captain or leader, in school.

d. I have lead a sports or games team as captain or leader outside school.

e. I have won a prize in sports or games as a member of a team.

f. I have won an individual prize in a sports or games competition.

ANY OTHER ACHIEVEMENT

APPENDIX C

Kelley's Formulae and Computer Programme

```
BEGIN
   %REAL M,SD
   %INTEGER N
   READ(N);READ(M);READ(SD)
   %BEGIN
       %REAL DI,DP,MTP,MLP,DIJ,DS,MSP,VTP,VLP,SDTP,SDLP,CDS,VSP,SDSP,QIJ
       %REALARRAY Q,Z,X(1:N)
       %INTEGER I,J
       %CYCLE I=1,1,N;READ(Q(I));READ(Z(I));READ(X(I));%REPEAT
       %PRINTTEXT'              PORTION MEAN          ';SPACES(35)
       %PRINTTEXT'SLICE MEAN';NEWLINE
       %PRINTTEXT'PERCENT      TOP        LOWER      '
       %CYCLE I=1,1,N;SPACES(2);PRINT((1-Q(I))*100,2,1);SPACES(3);%REPEAT
       NEWLINES(2)
       %CYCLE I=1,1,N;SPACE;PRINT(100*Q(I),2,1);SPACES(3)
       DI=Z(1)/Q(I)                                        d = z/q
       DP=SD*DI                                            d' = σz/q
       MTP=M+DP          Top portion mean = m+d'
       MLP=M-DP          Lower portion mean = m-d'
       PRINT(MTP,3,3);SPACES(2);PRINT(MLP,3,3);SPACES(3)
       %CYCLE J=1,1,N;QIJ=1-Q(I)-Q(J)
       DIJ=(Z(J)-Z(I))/QIJ                                 d_ij = (z_j - z_i)/(q_j - q_i)
       DS=SD*DIJ                                           d'_ij = σ d_ij
       MSP=M+DS          Slice mean = m + d'_ij
       PRINT(MSP,3,3);SPACES(2)
       %REPEAT;NEWLINE;%REPEAT;NEWLINES(3)
       %PRINTTEXT'           STANDARD DEVIATION    ';SPACES(28)
       %PRINTTEXT'SLICE STANDARD DEVIATION';NEWLINE
       %PRINTTEXT'PERCENT      TOP        LOWER      '
       %CYCLE I=1,1,N;SPACES(2);PRINT(100*(1-Q(I)),2,1);SPACES(3);%REPEAT
       NEWLINES(2)
       %CYCLE I=1,1,N;PRINT(100*Q(I),2,1);SPACES(3)
       DI=Z(I)/Q(I)
       VTP=(SD**2)*(1+X(I)*DI-DI**2)                       V' = σ²(1 + xz/q - d²)
       VLP=(SD**2)*(1+X(I)*DI-DI**2)
       SDTP=SQRT(VTP)                                      s.d. = √V'
       SDLP=SQRT(VLP)
       PRINT(SDTP,3,3);SPACES(2);PRINT(SDLP,3,3);SPACES(3)
       %CYCLE J=1,1,N;QIJ=1-Q(I)-Q(J)
       DIJ=(Z(J)-Z(1))/QIJ
       CDS=(X(J)*Z(J)+X(I)*Z(I))/QIJ
       VSP=(SD**2)*(1-CDS-DIJ**2)                          Slice V'_ji = σ²[1 + (x_j z_j - x_i z_i)/(q_j - q_i) - ((z_j - z_i)/(q_j - q_i))²]
       SDSP=SQRT(VSP)
       PRINT(SDSP,3,3);SPACES(2)               Slice s.d. = √(Slice V')
       %REPEAT;NEWLINE;%REPEAT;NEWLINES(3)
       %END
%ENDOFPROGRAM
```

	PORTION MEAN		SLICE MEAN					
PERCENT	TOP	LOWER	95.0	90.0	85.0	80.0	75.0	70.0
5.0	133.530	87.062	110.296	111.255	112.127	112.952	113.750	114.534
10.0	130.064	90.528	109.337	110.296	111.162	111.977	112.761	113.529
15.0	127.805	92.787	108.465	109.430	110.296	111.107	111.884	112.642
20.0	126.063	94.529	107.640	108.615	109.485	110.296	111.070	111.822
25.0	124.613	95.979	106.842	107.831	108.708	109.522	110.296	111.045
30.0	123.351	97.241	106.058	107.063	107.950	108.770	109.547	110.296
35.0	122.216	98.375	105.279	106.304	107.204	108.032	108.814	109.565
40.0	121.175	99.417	104.496	105.546	106.462	107.300	108.089	108.845
45.0	120.203	100.389	103.703	104.781	105.716	106.568	107.366	108.128

	STANDARD DEVIATION		SLICE STANDARD DEVIATION					
PERCENT	TOP	LOWER	95.0	90.0	85.0	80.0	75.0	70.0
5.0	4.185	4.185	8.891	8.189	7.635	7.160	6.734	6.341
10.0	4.633	4.633	8.189	7.452	6.870	6.370	5.921	5.506
15.0	4.973	4.973	7.635	6.870	6.267	5.750	5.285	4.855
20.0	5.267	5.267	7.160	6.370	5.750	5.218	4.741	4.300
25.0	5.538	5.538	6.734	5.921	5.285	4.741	4.255	3.805
30.0	5.793	5.793	6.341	5.506	4.855	4.300	3.805	3.348
35.0	6.043	6.043	5.971	5.112	4.448	3.883	3.380	2.917
40.0	6.290	6.290	5.615	4.734	4.055	3.481	2.971	2.503
45.0	6.538	6.538	5.267	4.362	3.671	3.088	2.571	2.098

Frequency Distribution of IQ

IQ	Freq.	Cum. %age	IQ	Freq.	Cum. %age
140	2	100.0	110	7	49.1
139	1	98.1	109	3	42.6
134	1	97.2	108	2	39.8
131	1	96.3	107	1	38.0
130	1	95.4	106	7	37.0
128	2	94.4	105	3	30.6
126	1	92.6	104	3	27.8
125	2	91.7	103	3	25.0
124	3	89.8	102	4	22.2
122	2	87.0	101	2	18.5
121	1	85.2	100	3	16.7
120	1	84.3	98	2	13.9
119	3	83.3	97	1	12.0
118	2	80.6	96	2	11.1
117	4	78.7	93	1	9.3
116	3	75.0	92	1	8.3
115	5	72.2	91	1	7.4
114	4	67.6	90	3	6.5
113	4	63.9	89	3	3.7
112	4	60.2	85	1	0.9
111	8	56.5	-	-	-

APPENDIX D

Review of Literature

NOTES FOR APPENDIX D

1.	DT	Divergent Thinking Tests
2.	R or TG	Intelligence measure taken from school records = R; test given as part of research = TG.
3.	Scoring	Divergent Thinking tests scored for FL = Fluency SF = Spontaneous Flexibility O = Originality E = Elaboration A = Adequacy
4.	F, M	F = Female, M = Male
5.	NFER,PV-3	National Foundation for Educational Research, Primary Verbal 3.
6.	-	No information for this column in the source.
7.	ACER	Australian Council for Education Research
8.	CTMM	California Test of Mental Maturity
9.	CMMI	California Mental Maturity Inventory
10.	STEP	Sequential Tests of Educational Progress
11.	L-T Scale (V+NV)	Large-Thorndike Intelligence Scale Varbal and Non-verbal
12.	NAAQ	Non-academic Accomplishments Questionnaire taken from Wallach and Wing (1969).
13.	PMA	Primary Mental Abilities Intelligence Test
14.	AH4	Alice Heim 4 Intelligence Test
15.	SAT-V	Scholastic Aptitude Test-Verbal
16.	VRT	Verbal Reasoning Test
17.	MDTB (V+P)	Morrisby Differential Test Battery (Verbal and Perceptual)
18.	DAT	Differential Aptitude Test

AUTHORS	SUBJECTS			INTELLIGENCE TEST USED				DIVERGENT THINKING TESTS USED				OTHER VARIABLES STUDIED	CORRELATION OF DT[1] WITH		SIGNIFICANT POSITIVE RELATIONSHIP OF OTHER VARIABLES WITH	
	No.	Sex	Age/ Grade	Test	mean/ range	s.d.	R or TG[2]		Scoring[3]	Verbal/ Non-V	Reliability		IQ	DT	IQ	DT
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
BENNETT, S.N. (1973)	331	MF[4]	10yrs 3mths	NFER[5] PV	104.8	17.0	TG	Ideational Fluency, Spontaneous Flexibility, Associative Fluency, Originality, Expressional Fluency	FL SF O	V	-[6]	English Attainment, Imaginative Story	.54F .58M	Aver .51F .47M	Eng Attainment .87F .90M	Eng Attainment .55F .57M
BIGGS, FITZGERALD + ATKINSON (1971)	174	MF	11 yrs	ACER[7] group verbal	106	-	R	Uses, Hidden Figures,	SF	V NV	-	Raven's Matrices, Maths Concept Test, Implied Meaning (ACER), Teachers' Ratings of: Distractibility, Asks Questions, Independence etc.	.211 to .354 Ave. r=.270	.114 to .455 Ave. r=.244	Conceptualism Asks Questions, Independence, Not Easily Distractable	Conceptualism Asks Questions, Independence, Likes Novelty
BENTLEY, J.C. (1966)	75	MF	Grad-Students	Miller Analogies Test	-	-	R	Improvements, Uses, Ask and Guess,	-	V	-	Tests of: Cognition, Memory, Convergence, Divergence, Evaluation	.12	-	Cognition, Memory, Evaluation, Total Achievement, New Idea-Convergent	Divergence, Evaluation, Total Achievement, New Idea-Divergent
BOERSMA + O'BRYAN (1968)	46	M	11 yrs	Lorge-Thorndike (V+NV)	106 to 111	11 to 13	TG	Figure Completion, Unusual Uses,	SF,F O,E	V NV	-	-	-.101 to .554	.340 to .736	-	-
CICIRELLI, V.G. (1965)	609	MF	6th grade	CTMM[8]	112	14	R	Ask and Guess, Prod. Improvement, Uses, Questions, Just Suppose, Picture Completion, Figure Completion, Parallel Lines	SF F O E	V NV	-	Calif. Arith. Test Calif Lang. Test Gates Basic Reading Test	.09 to .24	-	Reading r=.66 Arithmetic r=.67 Language r=.67	Reading r.13 to .37 Arithmetic r.11 to .26 Language r.11 to .31

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
CLARK, VELDMAN + THORPE (1965)	192	MF	11-15 years	CTMM	125	12	R	Consequences, Common Situations, Seeing Problems	O	V	.83	Word Fluency, Calif. Ach. Test (Reading), Holtzman Inkblots, Catell J. Pers. Quiz,	.04	-	Previous Years GPA	W. Fluency Reading Holtzman Ink- blots
CLINE, V.B., RICHARDS, J.M., + ABE, C. (1962)	161	MF	High Sch Seniors	CMMI[9]	98	12	R	Consequences, Word Association, Hidden Figures, Uses, Matchstick Problems	SF FL O	V NV	-	GPA Over previous three years	.32F .35M Ave r	.22F .21M Ave r	Multiple r .68F .69M with GPA	Multiple r .62F .65M with GPA
CLINE, V.B., RICHARDS, J.M., + NEEDHAM, W.E. (1963)	114	MF	High Sch Seniors	CMMI	98 F 101 M	12 F 12 M	R	Consequences, Word Association, Hidden Figures, Brick Uses, Matchstick Problems	SF FL O	V. NV		GPA in Science, STEP[10], Teacher Rating for Sc. Perf., No. of Sc. Courses Taken, Quest. on Involvement with Science	Ave r .33F .35M	Ave r .24F .21M	Multiple r GPA .74F .70M STEP .66F .69M Teacher .66F Rating, .55M No. Sc. .73F Courses, .32M Involvmt .36 with Sc. .44	Multiple r GPA .61F .63M STEP .47F .68M Teacher .59F Rating, .52M No. Sc. .48F Courses, .32M Involvmt .29F with Sc. .44M
CROPLEY, A.J. (1966)	320	MF	7th Grade	L-T Scale[11] V+NV	114	15	TG	Uses, Consequences, Seeing Problems, Symbol Production, Circles, Hidden Figures, Remote Associates	O	V NV	.33 to .53	Academic Average	Factor Scores Oblique Rotat- ion .514	-	Remote Associates r.394	-

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
CROPLEY, A.J. (1972)	111	MF	12th Grade	L-T*	119	14	R	Seeing Problems, Tin Can Uses, Circles, Consequences, Symbol Production, Hidden Figures	0	V NV	-	Achievement in: Art, Drama, Music + Literature (NAAQ)[12]	Ave r .244F .272M	Ave r .149F .139M	Ave r .089 with NAAQ	Ave r .056 with NAAQ
CROPLEY, A.J. + MASLANY (1969)	207	MF	20 yrs Univ. Students	PMA[13]	28 to 39	7 to 10	TG	Instances, Uses, Similarities, Pattern Meanings, Line Meanings	0	V NV	.67 to .85	-	.006 to .200 mdn r .078	.267 to .742 mdn r .441	-	-
DACEY, J. MADAUS, G. + ALLEN, A. (1969)	182	MF	13-16 yrs	AH4[14]	66	15	TG	Product Improvement, Unusual Uses, Figure Completion, Circles	-	V NV	-	-	.10	.18 to .94	-	-
DEWING, KATHLEEN (1970)	394	MF	7th Grade	ACER	114	11	R	Uses Circles Squares	FL 0	V NV	Ave r .390 to .688	Teacher Rating of Creativity, Peer Rating of Creativity, Leisure Interest, Creative Motivation, Imaginative Composition	.143	-	-	High Creative Performance on Teacher Rating, Peer Rating, Leisure Interest, Creative Motivation, Imaginative Composition
Di SCIPIO, W.J. (1970)	182	MF	19 yrs Univ Students	SAT-V[15]	-	-	R	Verbal Fluency	FL	V	.64 O .92 F	Eysenck Personality Inventory	.20 with Fluency Score	.67M .61F O+F score	-	Stable Extraversion
FLESCHER, I. (1963)	110	MF	6th Grade	CTMM	116 to 142	8 to 9	R	Word Association, Drawing, Uses, Plot Title, Composition	0 SF	V NV	-	General Anxiety, Test Anxiety, Metropolitan Achievement Tests (MAT)	.04 to .65 Ave r .09	-.19 to .36	School Attainment (MAT)	-

* IQ and DT scores taken from earlier study, Cropley (1966)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
HADDON, F.H. + LYTTON, H. (1968)	211	MF	11-12 yrs	Moray House VRT[16]	102	-	R	Circles, Shape of Dots, Block Printing, Uses for Shoe Box, Problems in Taking Bath, Imaginative Story	FL SF O E	V NV	-	School Atmosphere, Sociometric Status	.480	-	Higher r with Sociometric Status in Formal School	Informal School Atmosphere
HADDON, F.H. + LYTTON, H. (1971)	148	MF	15 yrs.	VRQ*	102	-	R	Circles, Shape of Dots, Block Printing, Uses for Shoe Box, Problems in Taking Bath, Imaginative Story	FL SF O E	V N	-	School Atmosphere, DT score and Age, Social Class, School Attainment, Career Choice/ Course, Teacher Rating for Creativity	.615 VRQ at 11 yrs	.621 DT at 11 + 15 yrs	English Maths SES	Teacher Rating for Creativity SES
HARGREAVES D.J. + BOLTON, N. (1972)	117	MF	10-11 yrs	MDTB[17] (V+P)	102	14	TG	Consequences, Uses, What Kind Is It, Picture Meaning, Stories, Picture Completion, Drawing, Word Meaning, Nonsense Words, Similarities	FL SF O E	V NV	-	Remote Associates, Personality Questionnaire, Images Test, Picture Preference	Ave r .45	Ave r .55	Remote Associates r .53	Remote Associates .43
HASAN, PARWEEN, + BUTCHER, H.J. (1966)	175	MF	13 yrs	Moray House VRT	102	12	R	Word Association, Uses, Fables, Arithmetic Problems, Word Fluency, Expressional Fluency, Remote Associates, Drawing, Circles	SF FL O	V NV	-	English Attainment, Arith. Attainment, Teachers Ratings for: Desirability as Pupil	.743 Ave r .42	.005 to .632 Ave r .25	Eng. Attain. Arith. Attain. Desirability as Pupil	Eng. Attain. Arith Attain Desirability as Pupil

*Verbal Reasoning Quotent used from earlier study, Haddon and Lytton (1968).

1	2	3	4	5	6	7	8	9	10
KOGAN, N, + PANKOVE, ETHEL (1972)	101	MF	16 yrs	DAT[18]	-	-	R	Uses, Similarities, Pattern Meanings, Line Meanings	FL O
LEITH, G. (1972)	106	MF	9, 11, 13 yrs	Raven's Matrices	-	-	TG	Uses, Associations, Similarities	FL O
LYTTON, H. + COTTON, A.C. (1969)	143	MF	14 yrs	-	111 + 112	- -	-	Circles, Shape of Dots, Block Printing, Uses for Shoe Box, Problems in Taking Bath, Imaginative Story	FL SF O E
VERNON, P.E. (1971)	400	MF	8th Grade	-	-	-	-	Circles, Patterns, Uses, Improvements, Similarities, Topics, Consequences	O
WARD, W.C., KOGAN, N. + PANKOVE ETHEL. (1972)	191	MF	5th Grade	Kuhlman* Anderson	92	12	R	Uses, Similarities, Pattern Meanings, Line Meanings	FL O
WODTKE, K.H. (1964)	105 to 183	MF	2nd to 5th Grade	L-T	-	-	TG	Ask and Guess, Product Improvement, Uses, Consequences, Picture Completion, Figure Completion, Circles	FL SF O E

1	11	12	13	14	15	16	17
KOGAN, N, + PANKOVE, ETHEL (1972)	V NV	-	NAAQ: Art, Leadership Social Service, Literature, Music, Dramatic Arts, Science	Ave r .188	DT/DT .670	NAAQ in small schools, Individual testing for DT (Age 11)	NAAQ in small schools, individual testing for DT (Age 15)
LEITH, G. (1972)	V	.6	Introversion-Extraversion, Anxiety, Moderate + Reduced Stress, Conditions MS, RS	Rank Corr -.119 to .329	Kendall's W .471 to .828	-	Under MS Introversion + Anxiety; Under RS Extraversion
LYTTON, H. + COTTON, A.C. (1969)	V NV	-	School Atmosphere, Social Class	.170	-	Social Class r .58	Social Class r .26
VERNON, P.E. (1971)	V NV	.758	Age, Multiple Vocabulary, Rorschach, Art Interests, Teacher Rating of Curiosity, Autobiography	.19 to .29	.316 to .338	-	Under Relaxed Condition Autobiog. marks for: Conceptual Maturity, Creativity
WARD, W.C., KOGAN, N. + PANKOVE ETHEL. (1972)	V NV	.94	Motivation: Immediate, Delayed + No Reward Treatment	Ave r .12FL .21,O	Ave r .59FL .22,O	-	Immediate + Delayed Reward Treatments
WODTKE, K.H. (1964)	V NV	.34 to .75	Luchins Water Jar Test Imaginative Writing	.36 to .46	-	Water Jar Test	Imaginative Writing

* IQ was available for only 2/3 of the sample from school records.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
YAMAMOTO, K. (1966)	461 + 827	MF	5th Grade	L-T Verbal	115 + 109	11 14	TG	Ask and Guess, Product Improvement, Unusual Uses, Circles	FL SF O E	V NV	.79 to .99	-	.33 and .39	- -	- -	- -
YAMAMOTO, K. + CHIMBIDIS, MARIA, E. (1966)	790	MF	5th Grade	L-T Verbal	110	14	TG	Ask and Guess, Improvements, Uses, Circles	FL SF O E	V NV	.79 to .99	Stanford Achievement Tests: Para/Word Meaning, Spelling, Language, Arith. Reasoning, Soc. Studies, Science Study Skills	.28	.81 to .95	Achievement Tests: r.49 to .65	Achievement Tests: r.15 to .31

REFERENCES

*Barron, F. and Welsh, G.S. (1952) Artistic perception as a possible factor in personality style: its measurement by a figure preference test. J. Psychol. 33, 199-203.

Anastasi, Anne (1968) Psychological Testing New York: Macmillan (3rd ed).

Anderson, C.C. (1966) A cognitive theory of the non-intellective correlates of originality. Behav. Sc. 11, 284-294.

Anderson, J.E. (1960) The nature of abilities. In Torrance, E.P. (ed) Education and Talent. Minneapolis:Univ. of Minnesota Press.

Argyle, M. and Robinson, P.W. (1962) Two origins of achievement motivation. Brit. J. Soc. clin. Psychol. 1, 107-120.

Arnold, Magda B. (1962) Story Sequence Analysis New York : Columbia Univ. Press.

Atkinson, J.W. (1958) Motives in Fantasy, Action and Society Princeton, N.J. : Van Nostrand.

Barron, F. (1955) The disposition toward originality. J. abnorm and soc. Psychol. 51, 478-485.

Barron, F. (1963) Creativity and Psychological Health Princeton : Van Nostrand.

Barron, F. (1969) Creative Person and Creative Process New York : Holt, Rinehart and Winston. *

Bennett, S.N. (1973) Divergent thinking abilities - a validation study. Brit. J. educ. Psychol. 43, 1-7.

Bentley, J.C. (1966) Creativity and academic achievement. J. educ. Res. 59, 269-272.

Bhavnani, Reena and Hutt, Corinne (1972) Divergent thinking in boys and girls. J. Child Psychol. Psychiat. 13, 121-127.

Biggs, J.B., Fitzgerald, D., and Atkinson, Sonia M. (1971) Convergent and Divergent abilities in children and teachers' ratings of competence and certain classroom behaviours. Brit. J. educ. Psychol. 41, 277-286.

Birney, R.C. (1958) Thematic content and the cue characteristics of pictures. In Atkinson, J.W. (ed) Motives in Fantasy, Action and Society Princeton, N.J. : Van Nostrand.

Boersma, F.J., and O'Bryan, K. (1968) An investigation of the relationship between creativity and intelligence under two conditions of testing. J. Person. 36, 341-348.

Bruckman, Idel R. (1966) Relationship between achievement motivation and sex, age, social class, school stream and intelligence. Brit. J. soc. clin. Psychol. 5, 211-220.

Burks, B.S. (1928) Genetic Studies of Genius. Vol. III. The Promise of Youth. Stanford Calif. : Stanford Univ. Press.

Burt, C. (1962) Critical notice of Creativity and Intelligence by Getzels and Jackson. Br. J. educ. Psychol. 32, 292-298.

Butcher, H.J. (1968) Human Intelligence : Its Nature and Assessment : London : Methuen.

Campbell, D.T. and Fiske, D.W. (1959) The convergent and discriminant validation by the multitrait-multimethod matrix. Psychol. Bull.56, 81-105.

Cartledge, C.J. and Krauser, E.L. (1963) Training first grade children in creative thinking under quantitative and qualitative motivation. J. educ. Psychol. 54, 295-299.

Cattell, R.B. and Butcher, H.J. (1968) The Prediction of Achievement and Creativity. Indianapolis : Bobbs-Merrill.

Chassell, Laura M. (1916) Tests for originality. J. educ. Psychol. 7, 317-328.

Child, I.L., Storm, T. and Veroff, J. (1958) Achievement themes in folk tales related to socialization practice. In Atkinson, J.W. (ed) Motives in Fantasy Action and Society. Princeton, N.J. : Van Nostrand.

Cicirelli, V.G. (1965) Form of the relationship between creativity, I.Q. and academic achievement. J. educ. Psychol. 56, 303-308.

Clark, C.M., Veldman, D.J. and Thorpe. J.S. (1965) Convergent and divergent thinking abilities of talented adolescents. J. educ. Psychol. 56, 157-163.

Classification of Occupations (1970) Office of Population Censuses and Surveys. London : HMSO.

Cline, V.B., Richards, J.M., Jr. and Abe, C. (1962) The validity of a battery of creativity tests in a high school sample. Educ. psychol. Measur. 22, 781-785.

Cline, V.B., Richards, J.M. Jr., and Needham,W.E. (1963) Creativity tests and achievement in high school science. J. appl. Psychol. 47, 184-189.

Coleman, J.S. (1960) Adolescent subculture and academic achievement. Amer. J. Soc. 65, 337-347.

Cowan, R.B. (1971) School children's Perception of and Preference for School Subjects M. Ed. Thesis. Univ. of Edinburgh.

Cox, Catherine M. (1926) Genetic Studies of Genius. Vol. II. The Early Mental Traits of Three Hundred Geniuses Stanford, Calif. : Stanford Univ. Press.

Crockett, H.J. (1962) The achievement motive and differential occupational mobility in the United States. Amer. Social. Rev. 27, 191-204

Cronbach, L.J. (1960) Essentials of Psychological Testing New York : Harper (2nd ed).

Cronbach, L.J. and Gleser, Goldine, C. (1965) Psychological Tests and Personnel Decisions Urbana : Univ. of Illinois Press.

Cropley, A.J. (1966) Creativity and intelligence. Brit. J. educ. Psychol. 36, 259-266.

Cropley, A.J. (1972) A five year longitudinal study of the validity of creativity tests. Develop. Psychol. 6, 119-124.

Dainton, F.S. (1968) Enquiry into the Flow of Candidates in Science and Technology into Higher Education London : HMSO

Datta, L.E. (1963) Test instructions and identification of creative scientific talent. Psychol. Rep. 13, 495-

de Charms, R., Morrison, H.W., Reitman, W.R. and McClelland, D.E. (1965) Behavioural correlates of directly and indirectly measured achievement motivation. In McClelland, D.C. (ed) Studies in Motivation New York:Appleton - Century Crofts.

de Charms, R. (1968) Personal Causation New York : Academic Press.

de Mille, R. and Merrifield, P.R. (1962) Book review of Creativity and Intelligence by Getzels and Jackson. Educ. and Psychol Measur. 22, 803-808.

Dellas, M. and Gaier, E.L. (1970) Identification of Creativity : the individual. Psychol. Bull. 73, 55-73.

Dentler, R.A. and Mackler, B. (1964) Originality : some social and personal determinants. Behav. Sc. 9, 1-7.

Di Scipio, W.J. (1971a) Divergent thinking and personality measures of English and American education majors. J. Genet. Psychol. 119, 99-107.

Di Scipio, W.J. (1971b) Divergent thinking : a complex function of interacting dimensions of extraversion - introversion and neureticism - stability. Brit. J. Psychol. 62, 545-550.

Douvan, Elizabeth (1956) Social status and success striving. J. abnorm. soc. Psychol. 52, 219-223.

Dunnette, M.D. (1963) A note on the criterion. J. Psychol. 47, 251-253.

Edwards, M.P. and Tyler, L.E. (1965) Intelligence, creativity and achievement in non-selective public junior high school. J. educ. Psychol. 56, 96-99.

Elkind, D., Deblinger, Joann and Adler, D. (1970) Motivation and creativity : the context effect. Amer. Educ. res. Jour. 7, 351-357.

Eysenck, H.J. (1967) Intelligence assessment:a theoretical and experimental approach. Brit. J. educ. Psychol. 37, 81-98.

Field, W.F. (1951) The effects on thematic apperception of certain experimentally aroused needs. Doctoral dissertation, University of Maryland, referred to in McClelland et al (1953). The Achievement Motive, New York : Appleton-Century Crofts.

Finlayson, D.S. (1972) Expressed motivation in relation to achievement motive, neuroticism and school success. Brit. J. educ. Psychol. 42, 65-70.

Flescher, I. (1963) Anxiety and achievement of intellectually gifted and creatively gifted children. J. Psychol. 56, 251-268.

France, N. (1964) The use of group tests of ability and attainment : a follow-up study from primary to secondary school. Brit. J. educ. Psychol. 34, 19-33.

French, Elizabeth (1958) Some characteristics of achievement motivation. J. exp. Psychol. 57, 306-309.

Frick, J.W., Guilford, J.P., Christensen, P.R. and Merrifield, P.R. (1959) A factor-analytic study of flexibility in thinking. Educ. Psychol Measur. 19, 469-496.

Garwood, Dorothy, (1964) Personality factors related to creativity in young scientists. J. abnorm. soc. Psychol. 68, 413-419.

Getzels, J.W. and Jackson, P.W. (1962) <u>Creativity and Intelligence : Explorations with Gifted Students.</u> New York : Wiley.

Getzels, J.W. and Jackson, P.W. (1963) The highly intelligent and the highly creative adolescent : a summary of some research findings:in Taylor, C.W. and Barron, F (eds) <u>Scientific Creativity : Its Recognition and Development</u> New York : Wiley.

Ginsburg, G.P. and Whittemore, R.G. (1968) Creativity and Verbal ability : a direct examination of their relationship. Brit. J. educ. Psychol. 38, 133-139.

Golann, S.E. (1962) The creativity motive. J. Pers. 30, 588-600.

Golann, S.E. (1963) Psychological study of creativity. Psychol. Bull. 60, 548-565.

Guilford, J.P. (1950) Creativity. Amer. Psychol. 5, 444-454.

Guilford, J.P. (1956) The structure of intellect. Psychol. Bull. 56, 267-293.

Guilford, J.P. (1959) Three faces of intellect. Amer. Psychol. 14, 469-479.

Guilford, J.P. (1965) <u>Fundamental Statistics in Psychology and Education</u> New York : McGraw-Hill.

Guilford, J.P. (1967) <u>TheNature of Human Intelligence</u> New York : McGraw-Hill.

Guilford, J.P. (1968) <u>Intelligence, Creativity and their Educational Implications</u> Diego, Calif: Knapp.

Haber, R.N. and Alpert, R. (1958) The role of situation and picture cues in projective measurement of the achievement motive. In Atkinson, J.W. (1958)*. <u>Motives in Fantasy, Action and Society</u> Princeton, N.J. : Van Nostrand.

Haddon, F.A. and Lytton, H. (1968) Teaching approach and divergent thinking ability. Brit. J. educ. Psychol. 38, 171-180.

Haddon, F.A. and Lytton, H. (1971) Primary education and divergent thinking abilities - four years on. Brit. J. educ. Psychol. 41, 136-147.

Hargreaves, D. (1967) <u>Social Relations in the Secondary School</u> London : Routledge and Kegan Paul.

*Heider, F. (1958) The Psychology of Interpersonal Relations
New York : Wiley.

Hargreaves, D.J. and Bolton, N. (1972) Selecting creativity tests for use in research. Brit. J. Psychol. 63, 451-462.

Harrison, R. (1940) Studies in the use and validity the TAT with mentally disordered patients. III. Validation by the method of blind analysis. Charact. and Pers. 9, 134-138.

Hasan, Parween (1965) Creativity and Intelligence B. Ed. thesis, University of Edinburgh.

Hasan, Parween and Butcher, H.J. (1966) Creativity and intelligence : a partial replication with Scottish children of Getzels and Jacksons' study. Br. J. Psychol. 57, 129-135.*

Heim, Alice (1970) Intelligence and Personality London : Penguin

Henry , W.E. (1956) The Analysis of Fantasy : The Thematic Apperceptive Technique in the Study of Personality. New York : Wiley.

Holland, J.L. (1961) Creative and academic performance among talented adolescents. J. educ. Psychol. 52, 136-147.

Holland, J.L. (1966) The prediction of academic and non-academic accomplishments in Proceedings of the 1966 Invitational Conference on Testing Problems Princeton, N.J.: Educational Testing Service.

Holland, J.L. and Astin, A.W. (1962) The prediction of the academic, artistic, scientific and social achievement of under-graduates of superior scholastic aptitude. J. educ. Psychol. 53, 132-143.

Holland, J.L. and Nichols, R.C. (1964) Prediction of academic and extra-curricular achievement in college J. educ. Psychol. 55, 55-65.

Holland, J.L. and Richards, J.M.Jr. (1965) Academic and non-academic accomplishments : correlated or uncorrelated: J. educ. Psychol. 56, 165-174.

Hoyt, D.P. (1965) The relationship between college grades and adult achievement. A review of the literature research reports No. 7. Iowa City : American College Testing Programme.

Hudson, L. (1964b) Academic sheep and research goats. New Society, 108 p.9.

Hudson, L. (1965) Intelligence Convergent and Divergent. in Barnett, S.A. and McLaren Ann (eds) Penguin Science Survey B London : Penjuin

Hudson, L. (1966) Contrary Imaginations London : Penguin.

Hudson, L. (1968) Frames of Mind London : Penguin.

Jackson, P.W. and Messick, S. (1965) The person, the product and the response : conceptual problems in the assessment of creativity. J. Pers. 33, 309-329.

Jacobs, B.jr. (1958) A method for investigating the cue characteristics of pictures. in Atkinson, J.W. (ed) Motives in Fantasy, Action and Society Princeton, N.J. : Van Nostrand.

Jayasuria, D.L. (1960) Study of adolescent ambition, level of aspiration and achievement motivation. Ph.D. thesis, University of London.

Kagan, J. and Moss, H.A. (1959) The Stability and validity of achievement fantasy. J. abnorm. soc. Psychol. 58, 357-364.

Katz, I. (1967) The socialization of academic motivation in minority group children. in Levine, D. (ed) Nebraska Symposium on Motivation Lincoln, Nebraska : Univ. of Nebraska Press.

Keddie, Nell (1971) Classroom knowledge. in Young, M.F.D. (ed) Knowledge and Control London : Collier - Macmillan.

Kelley, T.L. (1947) Fundamentals of Statistics Camb. Mass : Harvard Univ. Press.

Ketcham, W.A. and Kheiralla, S. (1962) Creativity in relation to intelligence and school achievement. In Collected Papers of the Inter-Institutional Seminar. Education Department, Henry Ford Museum.

Klein, S.P. and Evans, F.R. (1969) Early predictors of later creative achievements. Proceedings of the 77th Annual Convention, American Psychological Association, 153-154.

Knapp, R.H. (1963) Demographic, cultural and personality attributes of scientists. in Taylor, C.W. and Barron, F. (eds) Scientific Creativity : Its Recognition and Development New York : Wiley.

Kogan, N. and Morgan, F.T. (1969) Task and motivational influences on the assessment of creative and intellective ability in children. Genet. psychol. Monog. 80, 91-127.

Kogan, N. and Pankove, Ethel (1972) Creative ability over a five-year span. Child Dev. 43, 427-442.

Lacey, C. (1970) Hightown Grammar Manchester : Manchester Univ. Press.

Lavin, D.E. (1965) The Prediction of Academic Performance. New York : Russell Sage Foundation.

Leith, G. (1972) The relationship between intelligence, personality and creativity under two conditions of stress. Br. J. educ. Psychol. 42, 240-247.

Lindzey, G. and Heinemann, Shirley H. (1955) Thematic Apperception Test : individual and group administration. J. Pers. 24, 34-55.

Lowell, E.L. (1952) The need for achievement on learning and speed of performance. J. Psychol. 33, 31-40.

Lytton, H. and Cotton, A.C. (1969) Divergent thinking abilities in secondary schools Br. J. educ. Psychol. 39, 188-190.

Mackinnon, D.W. (1962a) The nature and nurture of creative talent. Amer. Psychol. 17, 484-495.

Mackinnon, D.W. (1962b) The personality correlates of creativity : a study of American architects. in Proceedings of the Fourteenth Congress on Applied Psychology, Vol. 2. Munksgaard.

Mackinnon, D.W. (1965) Personality and the realisation of creative potential. Amer. Psychol. 20, 273-281.

Madaus, G.F.A. (1967) A cross-cultural comparison of the factor structure of selected tests of divergent thinking. J. soc. Psychol. 73, 13-21.

Maddi, S.B. (1965) Motivational aspects of creativity. J. Pers. 33, 330-347.

Marino, C.J. (1971) Creativity and conformity among school children as influenced by religious application and type of school attended. Br. J. Soc. clin. Psychol. 10, 132-137.

Marsh, R.W. (1964) A statistical reanalysis of Getzels and Jacksons' data. Br. J. educ. Psychol. 34, 91-93.

*Miller, G.A. (1964) Book Review of A. Koestler's The Act of Creation
Scien. Amer. 211, p.144.

Maw, W.H. and Maw, E.W. (1970) Nature of creativity in high and low curiosity boys. Dev. Psychol. 2, 325-329.

McClelland, D.C. (1958) The use of measures of human motivation in the study of society. in Atkinson, J.W. (1958). op. cit.

McClelland, D.C. (1962) On the psychodynamics of creative physical scientists. in Gruber, H.E., Terrell, G. and Wertheimer, M. (eds) <u>Contemporary Approaches to Creative Thinking</u> New York : Atherton.

McClelland, D.C., Atkinson, J.W., Clark, R.A., and Lowell, E.L. (1953) <u>The Achievement Motive</u> New York : Appleton.

McNemar, Q. (1962) <u>Psychological Statistics</u> NewyYork : Wiley (3rd ed).

McNemar, Q. (1964) Lost : Our intelligence? Why? Amer. Psychol. 19, 871-882.

Mednick, S.A. (1962) The associative basis of the creative process. Psychol. Rev. 69, 220-232.

Mednick, S.A. and Mednick, Martha T., (1964) An associationist view of creative thinking in Taylor, C.W. (ed) <u>Widening Horizons of Creativity</u>, New York : Wiley. *

Murray, H.A. (1943) <u>Thematic Apperception Test Manual</u> Camb. Mass. Harvard Univ. Press.

Murstein, B.I. (1963) <u>Theory and Research in TAT</u> New York : Wiley.

Nicholls, J. (1970) Book review of <u>The Talented Student</u> by Wallach, M. and Wing, C.W. jr. in J. Gent. Psychol. 117, 277-279.

Nisbet, J.D. and Entwistle, N.J. (1969) <u>The Transition to Secondary Education</u> London : Univ. of London Press.

Ogletree, E. (1971) A cross-cultural examination of creative thinking ability of public and private school pupils in England, Scotland and Germany. J. Soc. Psychol. 83, 301-302.

Pankove, Ethel and Kogan, N. (1968) Creative ability and risk taking in elementary school children. J. Pers. 36, 420-439.

Parnes, S.J. and Meadow, S. (1959) Effects of"brain-storming" instructions on creative problem-solving by trained and untrained subjects. J. educ. Psychol. 50, 171-176.

Reich, C. (1970) The Greening of America London : Penguin.

Reitman, W.R. and Atkinson, J.W. (1958) Some methodological problems in the use of thematic apperceptive measures of human motives. in Atkinson, J.W. (1958).

Ricciuti, H.N. and Clark, (1957) A comparison of Need-Achievement Stories written by Experimentally "Relaxed" and "Achievement-Oriented" Subjects : Effects Obtained with New Pictures and Revised Scoring Categories. Tech. Report. Office of Naval Research Contract Nonr-694 (00), Princeton, N.J. : Educational Testing Service.

Ricciuti, H.N. and Shultz, D.G. (1958) Level of Aspiration Measures and Self-Estimates of Personality in Relation to Achievement Motivation, Tech. Report, Office of Naval Research Contract Nonr-694 (00), Princeton, N.J. Educational Testing Service.

Richards, J.M. jr., Holland, J.L. and Lutz, Sandra (1967) Prediction of student accomplishment in college J. educ. Psychol. 58, 343-355.

Robinson, P.W. (1964) The achievement motive, academic success and intelligence test score. Brit. J. soc. clin. Psychcl. 4, 98-103.

Rock, D.A., Evans, F.R. and Klein, S.P. (1969) Predicting Multiple Criteria of Creative Achievements with Moderator Variables. Research Bulletin RB-69-34, Princeton, N.J. : Educational Testing Service.

Roe, Anne (1952) A psychologist examines sixty-four eminent scientists. Amer. Psychol. 187, 21-25.

Roe, Anne (1953) A psychological study of eminent psychologists and anthropologists and a comparison with biological and physical scientists. Psychol. Monogr. 67, No. 352.

Rosen, B.C. (1956) The achievement syndrome a psycho-cultural dimension of social stratification. Amer. Soc. Rev. 21, 203-211

Saddaca, R., Clark, R.A. and Ricciuti, H.N. (1957) Content Analysis of Achievement Motivation Protocols : A Working Manual. Tech. Report, Office of Naval Research Contract Nour-694(00) Princeton, N.J. : Educational Testing Service.

Saddaca, R., Ricciuti, H.N. and Swanson, E.O. (1956) Content Analysis of Achievement Motivation Protocols : A Study of Scorer Agreement. Tech. Report, Office of Naval Research Contract Nonr-694(00), Princeton, N.J. : Educational Testing Service.

Sarason, S.B., Davidson, K.S., Lighthall, F.F., Waite, R.R. and Ruebush, B.K. (1960) <u>Anxiety in Elementary School Children</u> New York : Wiley.

Shaefer, C.E. and Anastasi, Anne (1968) A biographical inventory for identifying creativity in adolescent boys. J. Appl. Psychol. 52, 42-48.

Shapiro, R.J. (1966) The identification of creative research scientists. Psychol. Africana. 11, 99-132.

Skager, R.W., Shultz, C.R. and Klein, S.P. (1965) Quality and quantity of accomplishments as measures of creativity. J. educ. Psychol. 56, 31-39.

Spearman, C.E. (1923) The Nature of Intelligence and the Principles of Cognition. London : Macmillan.

Stein, M.I. (1956) A transactional approach to creativity in Taylor, C.W. (ed.) Research Conference on the Identificantion of Creative Scientific Talent Utah, Univ. of Utah Press.

Swift, D.F. (1966) Social class and achievement motivation. Educ. Res. 8, 83-95.

Swift, D.F. (1967) Family environment and 11+ success : some basic predictors. Brit. J. educ. Psychol. 37, 10-21.

Swift, D.F. (1968) Social class and educational adaptation in Butcher, H.J. (ed) Educational Research in Britain Vol. I London : Univ. of London Press.

Taylor, C.W. and Ellison, R.L. (1964) Predicting creative performance from multiple measures. in Taylor, C.W. (ed) Widening Horizons in Creativity New York : Wiley.

Taylor, H.C. and Russell, J.T. (1939) The relationship of validity coefficients to the practical effectiveness of tests in selection : discussion and tables. J. appl. Psychol. 48, 485-490.

Terman, L.M. (1925) Genetic Studies of Genius Vol. I. Mental and Physical Traits of a Thousand Gifted Children Stanford, Calif: Stanford Univ. Press.

Terman, L.M. and Merrill, Maud A. (1960) Stanford-Binet Intelligence Scale Boston : Hougheton Mifflin.

Terman, L.M. and Oden, Melita H. (1947) Genetic Studies of Genius Vol. IV The Gifted Child Grows Up Stanford, Calif. : Stanford Univ. Press.

Terman, L.M. and Oden, Melita H. (1959) Genetic Studies of Genius Vol. V The Gifted Group at Midlife Stanford, Calif. : Stanford Univ. Press.

Thorndike, R.L. (1962) Some methodological issues in the study of creativity. in Proceedings of the 1962 Invitational Conference on Testing Problems Princeton, N.J. Educational Testing Service.

Torrance, E.P. (1960) Educational achievement of the highly intelligent and the highly creative : eight partial replications of the Getzels-Jackson study. Bureau of Educ. Research, Univ. of Minnesota.

Torrance, E.P. (1962) Guiding Creative Talent Englewood Cliffs, N.J. Prentice-Hall.

Torrance, E.P. (1964) Education and creativity. in Taylor, C.W. (ed) Creativity : Progress and Potential New York : McGraw-Hill.

Torrance, E.P. (1965) Rewarding Creative Behaviour Englewood Cliffs, N.J. : Prentice-Hall.

Torrance, E.P. (1966a) Torrance Tests of Creative Thinking : Directions Manual and Sorting Guide Princeton, N.J. : Personnel Press.

Torrance, E.P. (1966b) Technical-norms Manual for the Torrance Tests of Creative Thinking : Research Edition Princeton, N.J. : Personnel Press.

Torrance, E.P., Tan, C.A. and Allman, T. (1970) Verbal originality and teacher behaviour : a predictive validity study. J. teach. Educ. 21, 335-341.

Vernon, P.E. (1964) Creativity and Intelligence. Educ. Res. 6, 163-169.

Vernon, P.E. (1965) Ability factors and environmental influences. Amer. Psychol. 20, 273-

Vernon, P.E. (1966) Development of current ideas about intelligence tests in Meade, J.E. and Parkes, A.S. (eds) Genetic and Environmental Factors in Human Ability Edinburgh : Oliver and Boyd.

Vernon, P.E. (1969) Intelligence and Cultural Environment London : Methuen.

Vernon, P.E. (1971) Effects of administration and scoring on divergent thinking tests. Brit. J. Educ. Psychol. 41, 245-257.

Veroff, J. (1950) in McClelland et al (1953) The Achievement Motive Appleton - Century Crofts.

Veroff, J. (1961) Thematic apperception in a nationwide sample survey. On Kagan, J. and Lesser, G.S. (eds) Contemporary Issues in Thematic Apperceptive Methods Springfield, Ill. : Charles C. Thomas.

Walker, D.J. (1967) Creativity and high school climate. in Gowan, J.C. Demos, G.O. and Torrance, E.P. (eds) Creativity Its Educational Implications New York : Wiley.

Walker, E.L., Atkinson, J.W., Veroff, J., Birney, R.C., Dember, W. and Moulton, R.W. (1958) The expression of fear related motivation in thematic apperception as a function of proximity to an atomic explosion. In Atkinson, J.W. (ed) Motives in Fantasy Action and Society. Princeton, N.J. : Van Nostrand.

Wallach, M. (1970) Creativity. in Mussen, P.H. (ed) Manual of Child Psychology New York : Wiley.

Wallach, M. and Kogan, N. (1965) Modes of Thinking in Young Children New York : Holt, Rinehart and Winston.

Wallach, M. and Wing, C.W. jr. (1969) The Talented Student : A Validation of the Creativity-Intelligence Distinction New York : Holt Rinehart and Winston.

Ward, W.C., Kogan, N. and Pankove, Ethel (1972) Incentive Effects in children's creativity. Child Dev. 43, 669-676.

Wechsler, D. (1958) The Measurement and Appraisal of Adult Intelligence Baltimore : Williams and Wilkins (4th ed).

Weiner, B. and Kukla, Andy (1970) An attributional analysis of achievement motivation. J. pers. soc. Psychol. 15, 1-20.

Weiner, B., Frieze, Irene, Kukla, Andy, Reed, Linda, Rest, S. and Rosenbaum, R.M. (1971) Perceiving the Causes of Success and Failure New York : General Learning Press.

Weiner, B., Heckhausen, H. and Meyer, W-U. (1972) Casual ascriptions and achievement behaviour. J. pers. soc. Psychol. 21, 239-248.

Werts, C.E. (1967) The many faces of intelligence. J. educ. Psychol. 58, 198-204.

*Williams, P. (1965) The structure and influence of adolescent peer groups. In Torrance, E.P. and Strom, R.D. (eds) Mental Health and Achievement New York : Wiley.

Wiseman, S. (1967) The effect of restriction of range on correlation coefficient. Brit. J. educ. Psychol. 37, 248-252.*

Wilson, R.C., Guilford, J.P., Christensen, P.R. and Lewis, D.J. (1954) A factor-analytic study of creative thinking abilities. Psychometrica, 19, 297-311.

Winterbottom, Marian R. (1958) The relation of need for achievement to learning experiences in independence and mastery. In Atkinson, J.W. (1958).

Wodtke, K.H. (1964) Some data on the reliability and validity of creativity tests. Educ. psychol. Measur. 24, 399-408.

Yamamoto, K. (1961) Educational achievement of highly creative students : threshold of intelligence. University of Minnesota, Bureau of Educational Research.

Yamamoto, K. (1964a) The role of creative thinking and intelligence in high school achievement. Psychol. Rep. 14, 783-789.

Yamamoto, K. (1964b) Threshold of intelligence in academic achievement of highly creative students. J. exper. Educ. 32, 401-405.

Yamamoto, K. (1964c) A further analysis of the role of creative thinking in high school achievement. J. Psychol. 58, 277-283.

Yamamoto, K. (1965a) Multiple achievement battery and repeated measurements : a postscript to three studies on creative thinking. Psychol. Rep. 16, 367-375.

*Yamamoto, K. (1965c) Creativity. - a blind man's report on the elephant. J. Counsel. Psychol. 12, 428-434.

Yamamoto, K. (1965b) Effects of restriction of range and test unreliability on correlation between measures of intelligence and creative thinking. Brit. J. educ. Psychol. 35, 300-305.*

Yamamoto, K. and Chimbidis, Maria E. (1966) Achievement, intelligence and creative thinking in fifth grade children : a correlational study. Merrill-Palmer Quarterly, 12, 233-241.

Yates, A. and Pidgeon, D.A, (1957) Admission to Grammar Schools London : Newnes.

(1970) Scottish Council for Research in Education, A Study of Fifteen-Year-Olds London : Univ. of London Press.